IMAGE PROCESSING IN THE HUMAN VISUAL SYSTEM

by

Brent Baxter

A dissertation submitted to the faculty of the University of Utah in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Electrical Engineering

University of Utah

December 1975

UNIVERSITY OF UTAH GRA

SUPERVISORY COMM

of a disserta

Brent

I have read this dissertation and have found it to be of
doctoral degree.

7-22-75
Date

Thomas G. Stockham Jr.
Chairman, Supervisory Committee

I have read this dissertation and have fo for a
doctoral degree.

7-22-75
Date

Richard W. Grow
Member, Supervisory Committee

I have read this dissertation and have found it to be of satisfactory quality for a
doctoral degree.

7-22-75
Date

Lamar K Timothy
Member, Supervisory Committee

I have read this dissertation and have fou
doctoral degree.

July 22 1975
Date

Dietrich K. Gehmlich
Member, Supervisory Committee

I have read this dissertation and have a
doctoral degree.

22 July 1975
Date

Richard F. Riesenfeld
Member, Supervisory Committee

UNIVERSITY OF UTAH GRADUATE SCHOOL

FINAL READING APPROVAL

To the Graduate Council of the University of Utah:

I have read the dissertation of Brent Scott Baxter in its final form and have found that (1) its format, citations, and bibliographic style are consistent and acceptable; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the Supervisory Committee and is ready for submission to the Graduate School.

10 Sept. 75
Date

Richard F. Riesenfeld
Member, Supervisory Committee

Aj

Richard W. Grow
Chairman/Dean

Approved for the Graduate Council

Sterling M. McMurrin
Dean of the Graduate School

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENT

Appreciation is expressed to Dr. T.G. Stockham who supervised this work and whose standards of professional excellence are well known. The photographs reproduced here are the work of Mike Milochik and the software used in the experiments associated with this research was contributed to by many people including Raphael Rom, William Greer, George Randall and Douglas Greer. Special thanks are due Mike Callahan for numerous stimulating discussions.

In any work of this sort, encouragement from home is most valuable, and without the patience and help of my wife, Karen, it is doubtful that this study would have been completed.

ABSTRACT

This work extends the multiplicative visual model to include image texture as suggested by experiments [Campbell, Weisel] linking a low resolution Fourier analysis with neurons in certain parts of the visual cortex. The new model takes image texture into account in the sense that weak texture is accentuated and strong, high contrast texture is attenuated. This model is then used as the basis for an improved image enhancement scheme and an unusually successful method for restoring blurred images. In addition, it is suggested how the model may provide new insights into the problem of finding a quantitatively correct image fidelity criterion.

The structure of this model is described in relation to visual neurophysiology and examples are presented of images processed by the new techniques.

The research described here also shows how the retinex [Land] can be implemented in a new way which allows the required computations to be carried out on a rectangualr grid.

INTRODUCTION

The work described here draws from such diverse disciplines as opthalmology, neurophysiology, psychology and engineering to construct a model that shows how images are processed as they are passed from the retina to higher centers in the visual cortex. The objective of this research is to develop image processing techniques of practical importance that are compatible with, and take advantage of, properties of the visual system. To introduce the new results described in chapters 3 and 4, a collection of optical illusions is presented which illustrates some important properties of the human visual system. The stylized nature of the images used to produce these illusions makes it possible to excite certain visual mechanisms without distractions imposed by the complexity of naturally occurring images. Bright and dark stripes like those in figure 1 are observable almost anywhere if one takes the trouble to look for them but they usually go unnoticed because our attention is focused on other things. Insights gained by studying these patterns are used to form a new model of vision upon which a variety of practical image processing schemes are based.

Material in succeeding chapters refers to the illusions and the reader is encouraged to study each one until the effect it illustrates is clearly understood. Approximately a minute is required for the adaptation effects shown in figures 4 and 5 to develop fully and the necessary time must be invested to appreciate the phenomena they represent.

Chapter 2 begins with a brief review of Stockham's multiplicative

model of vision [1] and the theory of generalized linearity [2] upon which the model is founded. An example of image enhancement based on the model is presented which should help provide some historical perspective for the new research results in chapters 3 and 4. Except for some engineering details in section 2.5, this chapter is a review of earlier work.

In chapter 3 a new model of the visual system is described which incorporates some recently discovered associations [3] between two dimensional Fourier analysis and neural structures in the visual cortex. Accompanying this discussion are examples of several important new image processing techniques based on this model.

The following new results are described in chapter 3:

(1) A frequency selective model of vision is presented that has an important texture * equalization feature. The implications of this feature are explored and the means of implementing such a model are outlined.

(2) Image enhancement based on the new model is shown to be capable of emphasizing weak detail while softening strong, high contrast detail. Some kinds of images can be made more "readable" in the sense that weak textural detail, which may convey important information, is placed more in evidence by the enhancement operation.

(3) A greatly improved method for removing photographically induced

---

* Texture refers to the semi-random variations in light intensity by which a pile of leaves might be distinguished from a smooth surface painted a similar color. Strong or high contrast texture refers to large amplitude variations in light intensity and weak texture refers to small ones

blur is discussed in relation to the new model. The echo-like artifacts associated with earlier methods [4,5] are largely eliminated and the usual problems introduced by film grain noise and truncation of the blurred image are dealt with successfully.

(4) It is shown how the similarities between mechansims responsible for items 2 and 3 above lead to the notion that human vision has a limited capacity for restoring blurred images. The extent to which this deblurring can occur naturally is discussed.

These four new developments are significant because they associate image texture with properties of human vision in a direct way and they show how systems may be designed which can manipulate texture.

Chapter 4 deals with ratios of reflected light in the style Land's retinex [6,7]. At the beginning of the chapter an original method for implementing the retinex is described followed by an illustration of some interesting properties of the retinex. One is a color constancy property and another is the correct prediction of the Cornsweet illusion.

In the concluding chapter some ideas are presented on how a more accurate image fidelity criterion might be developed based on the texture equalization property of section 3.3. Some design considerations relating to the blur removal scheme are outlined which could lead to improved performance and a better understanding of the blur restoration process.

CHAPTER 1

OPTICAL ILLUSIONS

1.1 Introduction

The reason we find optical illusions so interesting and entertaining is that they often reveal facets of our cognitive processes which tend to to be overlooked. The value in these demonstrations lies in our ability to infer from them how visual information might be processed by the complex optical and neural mechanisms upon which vision depends. These illusions represent the work of others and their inclusion here is intended to help illustrate the new work presented in chapters 3 and 4.

Perhaps the most widely studied of the illusions in this chapter is the Mach band pattern in figure 1. An apparent bright band at the left of the light/dark transition is not present in the light reflected from the photograph in figure 1 but is induced in our perceptual mechanism by lateral inhibitory interconnections between certain information carrying cells in the retina [8,9]. It was pointed out by Mach [10] that a second difference operation performed in an early part of the visual pathway would account qualitatively for the bands.

Most people have the ablilty to distinguish between parallel and nonparallel lines. We live in a world liberally supplied with geometric forms containing straight, parallel lines and right angles and it is not too surprising to find that we can tell whether lines

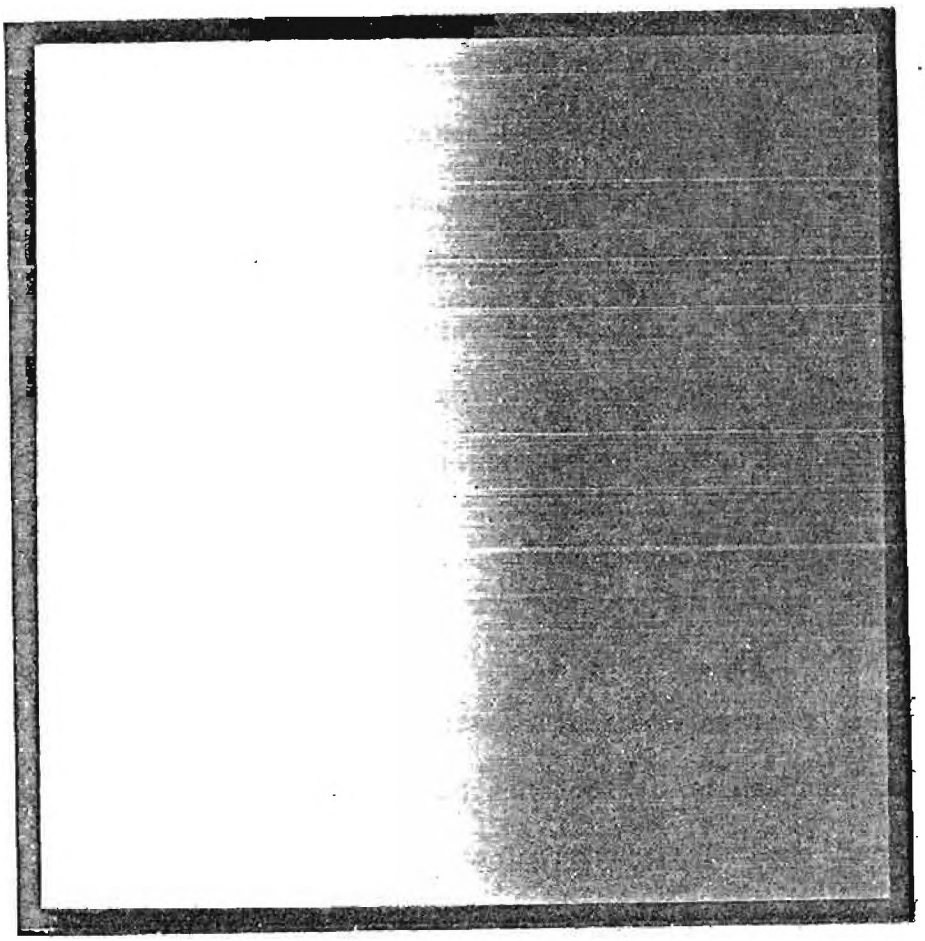
are drawn parallel or not. It is quite easy to undermine this ability by adding some appropriate distractions of the type shown in figure 6. It is obvious from the figure that this effect can be quite striking.

This chapter is devoted to describing seven optical illusions. The purpose is twofold: first, such a demonstration provides ample evidence that significant transformations take place as images are analyzed by the visual system, and second, each of the demonstrations illustrates a single aspect of vision as it influences the formulation of an accurate model of the visual system. Insights gained from these models will be used in developing new methods for processing images.

1.2 Mach Bands

The image in figure 1 was prepared under carefully controlled conditions to insure uniform reflectance throughout right and left patches, and linearly decreasing reflectance between them. Light and dark bands like the ones at the ends of the uniform regions may be seen almost anywhere a shadow is cast by an extended light source. A good example is the shadow cast on a wall by the sun shining through an open window. A dark band is usually visible in the dark part of the pattern where the sun's disk is just obscured by the window frame, and a light one is almost always visible on the bright side of the gradient.

A linear, shift invariant, spatial highpass filter produces artifacts qualitatively similar to Mach bands in response to an intensity gradient and it is one of the major constituents of the multiplicative model to be described in Chapter 2.



Mach Bands

Figure 1 -- Mach band phenomenon: Note the appearance of light and dark bands on opposite ends of the intensity gradient.

1.3 Simultaneous Contrast

Figure 2 shows a sinusoidal pattern surrounding two identical bars, which appear to be different shades of gray. When seeing this pattern for the first time, many people find it remarkable that

masking off the background with a template will cause both bars to appear identical in gray level. The term simultaneous is perhaps unnecessary since, in any image described as having contrast the contrasting elements occur simultaneously rather than sequentially, but the usage is common in the literature and it will be continued in this document. In Chapter 2 the implications of this illusion will be considered with regard to the multiplicative model of vision.

1.4 Herman Grid

A somewhat more subtle effect is shown in figure 3. At first glance it might seem to contain nothing more than a series of square patches separated by wide lines, but after looking steadily at one of the intersections for a few seconds the other intersections will seem to contain fuzzy spots which appear lighter (3a) or darker (3b) than the lines. It is not necessary for the patches to have abrupt boundaries *, although the effect is more striking if they do. A similar comment about abrupt boundaries applies to figure 2.

1.5 Adaptation

The term adaptation refers to a reduction in sensitivity to a stimulus by prolonged exposure to it. For example, when a person leaves a darkened movie theater and goes out into the bright sunlight his sensitivity to light is gradually reduced until the light level seems comfortable again.

---

* See Baudelaire [11] for a detailed study of contrast phenomena involving patterns of the type shown in figures 2 and 3 but with smooth boundaries.

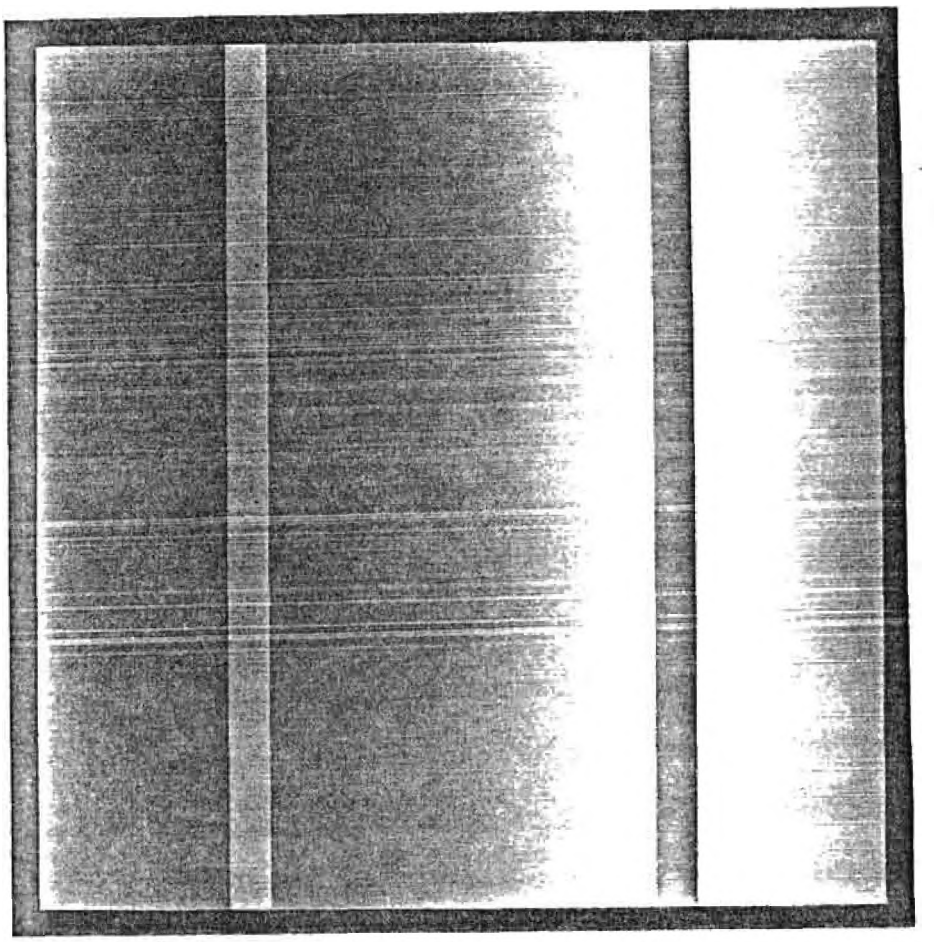
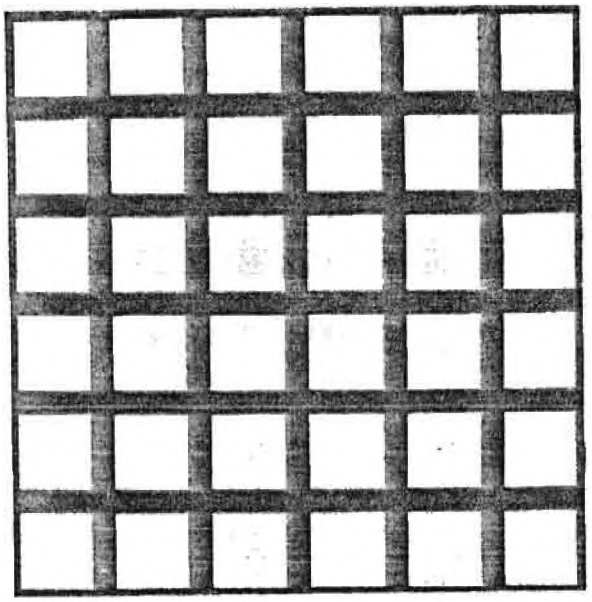
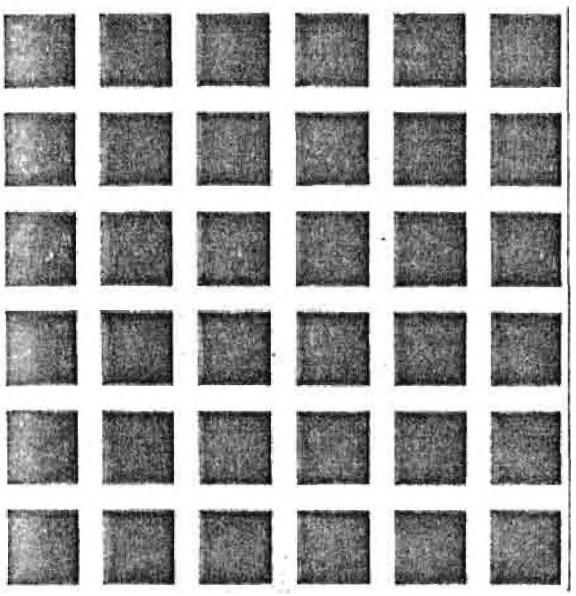


Figure 2 -- Simultaneous Contrast Phenomena: The two bars appear to be different shades of gray, even though they are really identical. This may be demonstrated by covering the sinusoidal background with the template provided in the pocket inside the rear cover.

The images of figure 4 form the basis for an adaptation experiment where the stimulus is not light level but the level of contrast in a bar pattern. Adaptation is caused by viewing a high contrast bar pattern and evidence of reduced sensitivity is provided by viewing a second bar pattern containing larger (3X) bars having lower contrast and noting a change in their appearance.



(a)



(b)

Figure 3 -- Herman Grid: Most observers report lighter (3a) and darker (3b) spots at the grid intersections after fixating some part of the grid for for a few seconds.

Central to the demonstration is a finding by Campbell et al. [12] with regard to the subjective appearance of low contrast bar patterns. He found that in order to be detectable, the pattern must have sufficient contrast for its fundamental component to be above the threshold of visibility; and for the bars to have well defined edges the third harmonic must also be above the visual threshold. This demonstration depends on changing the contrast sensitivity and therefore the visual threshold for third and higher * harmonics by adaptation. Evidence of this will be a change in appearance of figure 4(b).

Afterimage formation may interfere with the demonstration if a stationary image is allowed to remain on the retina. This can be prevented by permitting the center of attention to wander around the small circle during adaptation. After adapting to figure (4a) for about 60 seconds the bars in figure (4b) should appear blurred temporarily. Both steps are designed to be carried out with the page held at arms length.

As an interesting variation, the page might be rotated through 90 degrees between steps one and two. In this case there is no change in the appearance of the large bars indicating that adaptation of the kind discussed here raises the threshold only for patterns aligned with the adapting pattern. Another variation might be to

---

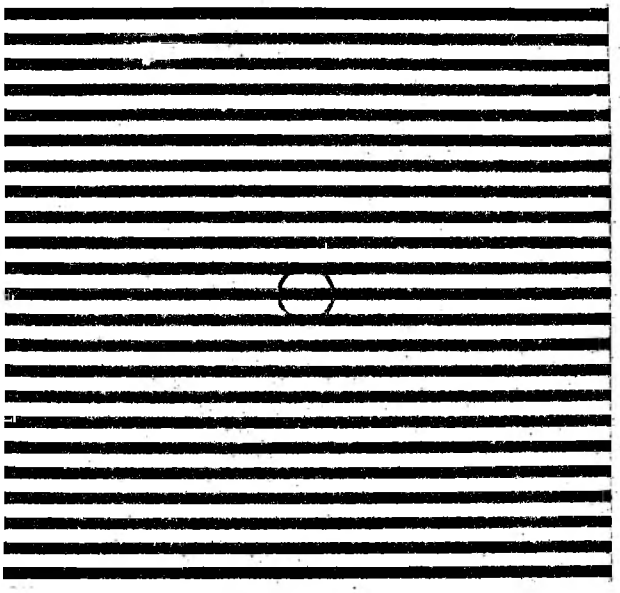
* Since the low contrast bars are three times as large as the bars in the adapting pattern, a reduction in sensitivity occurs only for the third, ninth, fifteenth ... harmonics of the large bars.

make the low contrast bars the same size as those in the adapting pattern. The low contrast bars will then disappear rather than become blurred.

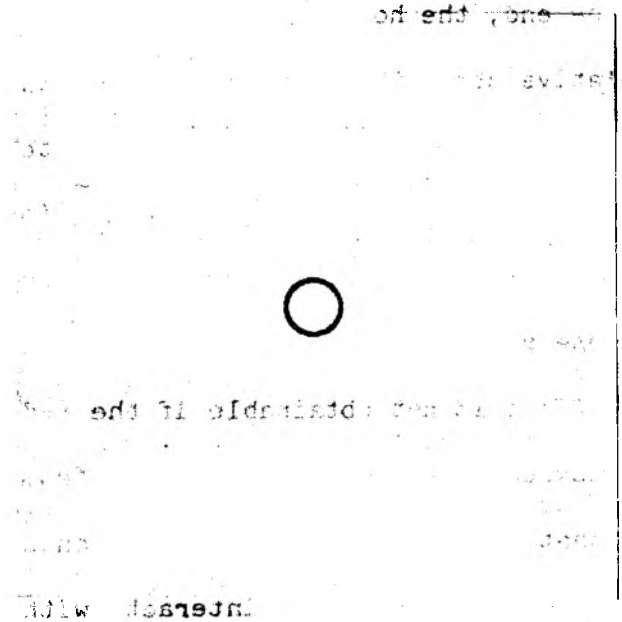
1.6 Monocular Alternation

A second demonstration suggesting spatial frequency adaptation, termed monocular alternation by Campbell [13], is shown in figure 5. Alternation refers to the fact that the horizontal (green) and vertical (red) bars seem to alternate in relative strength, and monocular indicates that the effect does not depend on binocular vision.

The reader should spend at least a minute viewing the pattern without interruption. Toward the end, the horizontal and vertical bars will seem to alternate in relative strength at about 2-5 second intervals. The time required for the effect to begin might indicate that some kind of electrochemical adaptation mechanism is responsible. The effect obtainable with a black and white image is not as strong as with the colored one suggesting that color may play a role, however the alternation effect is not obtainable if the red and green bars are parallel. An interesting suggestion arising from this effect is the possibility that adaptation might be based on a system of angularly sensitive filters which somehow interact with each other.



(a)



(b)

Figure 4 -- Viewing the high contrast bars (4a) for 30-60 seconds causes a reduction in sensitivity to horizontal bar patterns with similar spatial frequency content. Sensitivity to the third harmonic of the larger bars (4b) will be reduced enough to cause them to appear blurred temporarily.

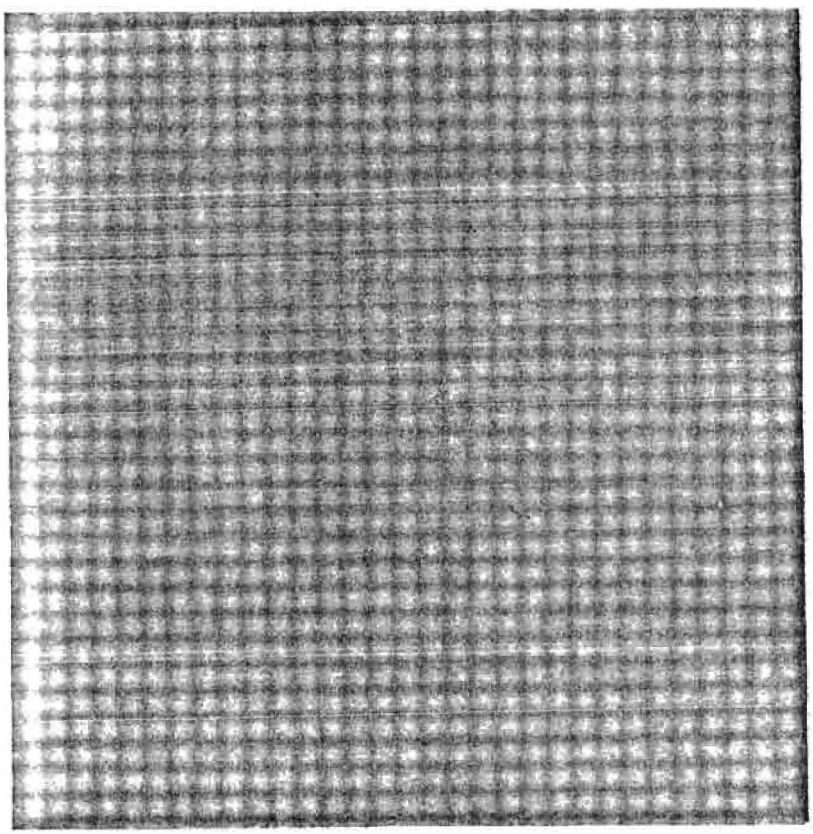


Figure 5 -- After viewing the pattern for about 60 seconds the horizontal and vertical bars should seem to alternate in relative strength. This demonstration is cited as evidence supporting the existence of angularly selective channels in the visual system.

1.7 Angular Illusions

Figure 6 is one of many geometric illusions that appear regularly in places such as the mathematical games section of Scientific American. They are usually accompanied by a comment like the following: "Parallel lines may be made to appear nonparallel when superimposed on a diagonal pattern." Such a superficial treatment of a striking effect might cause one to wonder why a more satisfying

not the long lines are straight and parallel, the reader is encouraged to check them with a straight edge and ruler. For most people the lines not only seem slightly curved and definitely nonparallel, but also seem to move slightly. This effect is apparent immediately and does not require prolonged viewing as did figures 4 and 5.

This is an interesting illusion for us to study because the effect is predicted correctly by the frequency selective model of chapter 3.

1.8 Cornsweet Illusion

An interesting edge induced effect named after T.N. Cornsweet [14] is shown in figure 7. Approaching from the left in figure (7a), there is an abrupt decrease in gray level followed by a gradual return to the original level. Figure (7b) is similar except for the circular symmetry. To most people the left half of (7a) appears darker than the right side. There is an important difference between these patterns and the ones in figures 2 and 3 in that the effect requires an abrupt change in gray level while the others occur without one although not as strongly. This distinction will be discussed later in connection with the retinex model of vision.

1.9 Summary

In this chapter several optical illusions have been introduced

which illustrate properties of the visual system important to our study of brightness

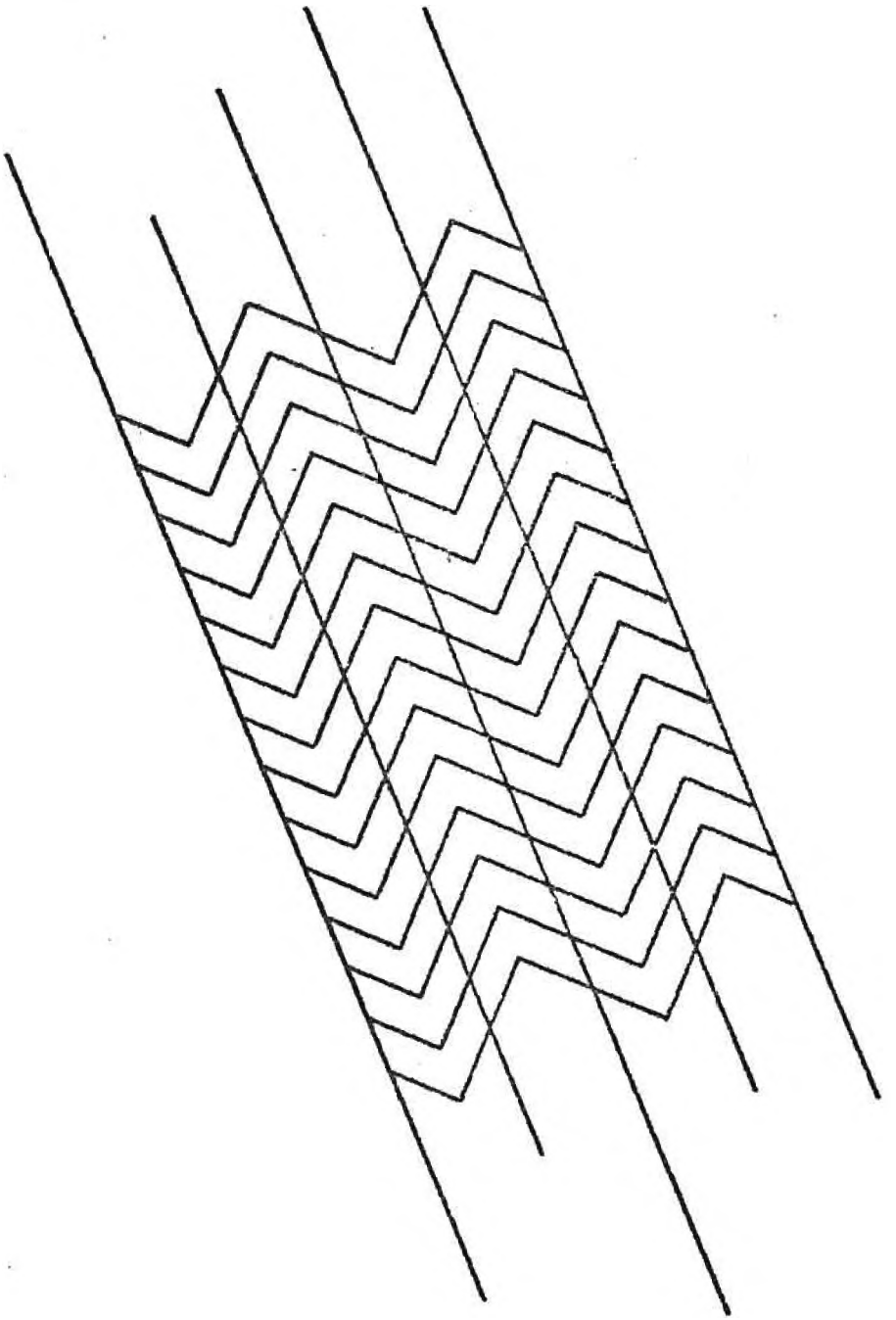
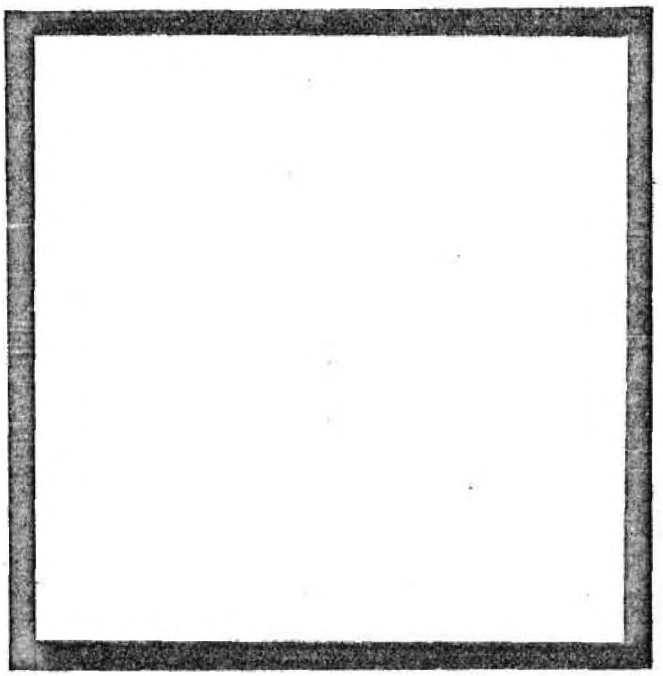


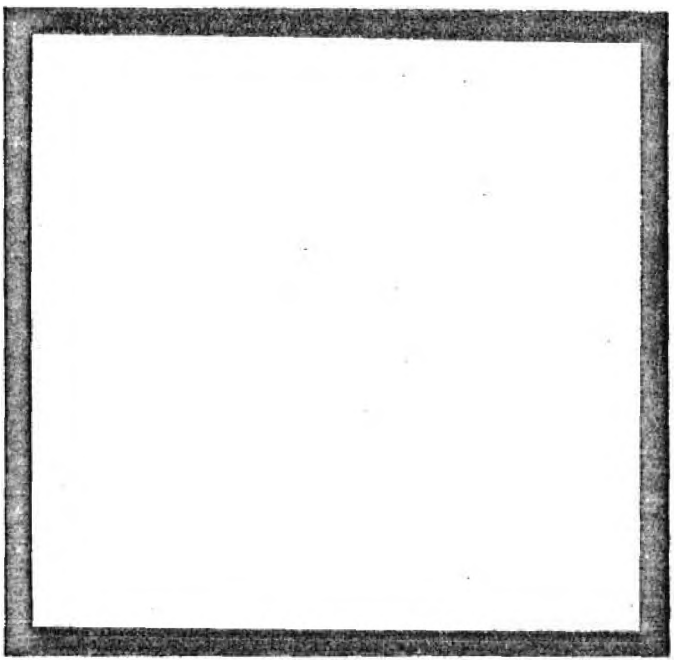
Figure 6 -- Angular illusion: The long vertical bars appear distinctly nonparallel due to influences exerted by the diagonal lines.

perception *. Mach bands are historically the oldest of these illusions and were probably the first indication of a spatial filtering mechanism associated with vision. Others such as the parallel line illusion and the Cornsweet illusion suggest more sophisticated processes. In chapters 2,3 and 4 implications of these demonstrations will be explored in an effort to formulate models of the human visual system upon which computationally realizable image processing methods can be based.

* The term brightness refers to the perceptual sensation evoked in the visual system by light. The relationship between two dimensional light patterns and brightness is the subject of this study.



(a)



(b)

Figure 7 -- In these examples of the Cornsweet illusion the left side (7a) appears darker than the right side even though they are an identical shade of gray at points a short distance from the edge.

CHAPTER 2

THE MULTIPLICATIVE MODEL

2.1 Introduction

Material in this chapter summarizes previous work by Stockham and others; and except for some experimental details in section 2.5, it is not original with the author. A review of generalized linearity as it applies to the multiplicative model [1,15,16] serves as an introduction to several new results presented in chapter 3. The utility of mapping multiplication into addition via the logarithm will be considered in connection with the multiplicative structure of image formation, and the use of linear filtering to remove illumination gradients will be discussed. Finally, the model's performance and limitations will be considered relative to the demonstrations of chapter 1.

2.2 Generalized Linearity

In a classic paper by Oppenheim et al. [2] the notion of mapping signals into a new space where they are combined by addition is shown to be part of a very powerful nonlinear filtering technique capable of separating signals combined by operations other than addition. In the case of multiplied signals, both of which are positive, the mapping consists of taking the logarithm. Other nonlinear filtering mechanisms such as the AGC units associated with the frequency

selective model and the deblurring scheme to be discussed later require more complicated transformations. A separation of the transformed signals may be performed by linear filtering provided their Fourier transforms do not overlap i.e., the transforms have disjoint support. The final step in this nonlinear filtering process is to map the signal of interest back into the original space by the inverse of the initial transformation.

The multiplicative model of brightness perception has a similar structure except that the final inverse transformation does not occur. The initial transformation is the logarithm which means that the model is capable of separating multiplied signals. Since the light reflected from an object consists of the illumination multiplied by surface reflectance, a mechanism is provided for removing illumination effects.

2.3 Logarithmic Encoding

One of the oldest observations suggesting that visual response might be related logarithmically to visual stimulus intensity resulted in what is known as Weber's Law [17,18]. Imagine a small test patch on a uniform background. If the patch is adjusted to be just noticeably lighter or darker than its background, Weber's Law implies that the amount by which the patch must be lighter or darker will be proportional to the background intensity. The constant of proportionality is known as the Weber fraction. The surprising thing about this relationship is the large range (about four orders of magnitude in light intensity) over which it is useful.

A credible hypothesis for explaining Weber's Law might include a logarithmic transformation followed by a mechanism sensitive to differences. Later investigators, notably Hartline [8,9], provided substantial electrophysiological evidence that neural activity in retinal cells of Limulus (the horseshoe crab) is approximately a logarithmic function of stimulus intensity, and similar results have been obtained for primates including humans [19]. Figure 9 shows this transformation explicitly.

It is fortunate that a mechanism like this is part of the visual process since it tends to weight detail in light and dark parts of a scene equally thus considerably extending the range of light intensity over which vision operates successfully. Light to dark ratios far in excess of 100:1 are common under natural outdoor lighting and without this equalization feature our ability to see detail over a wide range of light intensity would be seriously impaired. An ordinary black and white photographic negative is able to record wide intensity range images because of a logarithmic mechanism. The D Log E curve familiar to photographers, expresses a quasi-linear relationship between the amount of silver per unit area (density) fixed in a processed negative and the logarithm of exposure (intensity multiplied by exposure time). Light ratios as high as 10,000:1 can be recoded by a properly exposed and developed negative. Partly because of this logarithmic aspect of vision, Stockham [15] was able to make a persuasive case for employing a logarithmic representation in the transmission, processing and storage of images.

It is interesting to speculate on the effect these ideas might

have had on our commercial television system had they been understood at the time broadcasting standards were developed. A simpler television receiver might have emerged where adjustments in brightness could be made independently of changes in contrast. This is because adding a constant to the signal would have made the image lighter or darker without changing the contrast and scaling the signal would have changed only the contrast. The commercial system uses instead a representation of the form:

$$R = I^{\alpha} \quad \text{where,} \quad 0 < \alpha < 1$$

R = Representation
I = Image Intensity

This function is concave downward as is the logarithm and adjustments to brightness and contrast are made as described above but they interact. The reason for this representation is to correct for nonlinearities in the receiver CRT.

Another way in which the logarithmic representation of images is useful to vision has to do with our tendency to describe objects as light or dark depending on their light reflecting characteristics rather than on the total amount of light reflected from them. A sheet of white paper might reflect as much as 90% of the incident light and a man's black wool suit absorbs nearly all of the incident light, reflecting perhaps as little as 1%. It is possible to imagine situations, with the suit strongly illuminated and the paper weakly lighted and both visible at the same time, where more light is reflected from the suit than from the paper. Even under such unlikely conditions the suit will probably be perceived as dark and the paper will be perceived as light. The visual stimulus is greatly

modified in this situation by the unusual illumination, but perceptual judgments about light and dark objects remain relatively unchanged. In the next section it will be shown how linear filtering on a logarithmic representation of an image is capable of suppressing illumination effects.

2.4 Linear Filtering

The linear filtering mentioned in section 2.2 has a physiological counterpart in the retina in the form of lateral inhibitory connections between adjacent photoreceptors. In the stylized illustration of figure 8 these inhibitory connections are shown emanating from bipolar cells (white) and forming inhibitory junctions (green) where they attach to other bipolar cells. Hartline et al. [10] showed, that for a given retinal cell in Limulus, the inhibitory influences combine to give a central "on" region, where stimulation with light causes an increase in neural activity, surrounded by an anular "off" region, where activity is decreased by stimulation with light. These inhibitory effects have been shown to be linear * and are modeled by the linear filter in figure 9. It is interesting to note that the impulse response of a circularly symmetric highpass filter has a central peak surrounded by a region of opposite sign.

---

* In the absence of intensity steps, retinal interaction is approximately linear. For more details about the conditions under which linearity holds, see Baudelaire [chapt. 5 in 11].

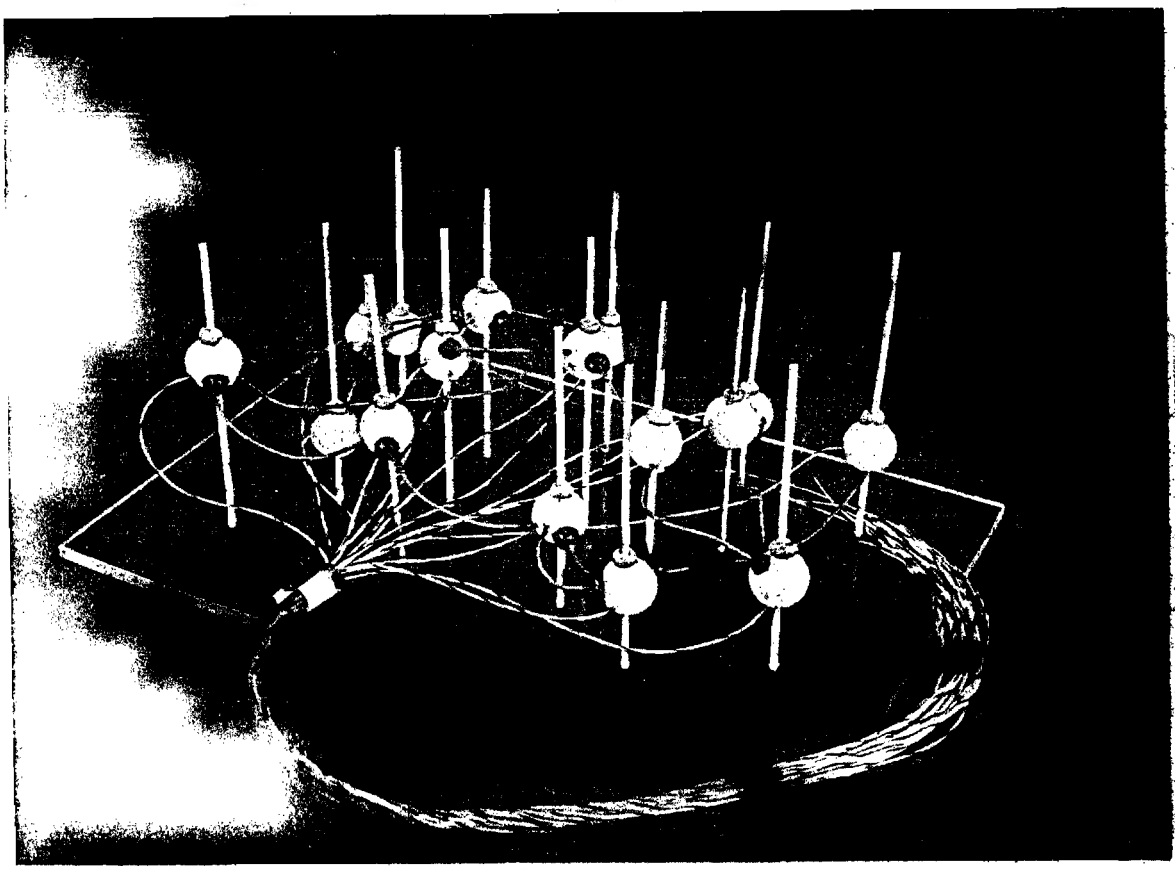


Figure 8 -- A physical model of the retinal cells and associated neural pathways. Inhibitory connections between bipolar cells are often considered to be responsible for Mach bands.

Legend: Red --------- Photoreceptors
White ------- Bipolar cells
Yellow clay - Excitatory connections
Green clay -- Inhibitory connections
Yellow wire - Optic nerve

Previously we have noticed the association between Mach bands and a highpass filter. A highpass filter will attenuate the sinusoidal background in figure 2 while partially preserving the large change in gray level between background and the bars. This helps explain the appearance of the figure.

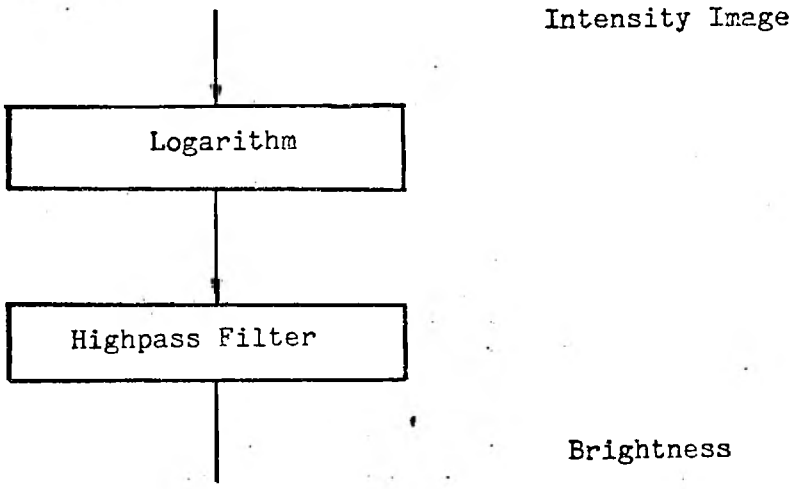


Figure 9 -- A Multiplicative model of brightness perception accounts for the eye's logarithmic response to light.

2.5 Image Enhancement

Whenever models are discussed the question always arises as to whether or not they have any practical utility. The light/dark compression feature of the logarithm has been successfully [15] used in connection with a linear highpass filter for image enhancement purposes [1,2]. The success of such a scheme is related to the structure of image formation: first, images are formed by reflection and consist of the product of illumination and reflectance

information; second, the logarithm of illumination is often a smooth function while the reflectance component is considerably less smooth. This relative difference in smoothness makes possible a partial separation using linear filtering. An example of the smoothness refered to above may be observed by noting how illumination changes smoothly in a typical room from the center of a wall to a corner.

Implementation of an image enhancement scheme requires a properly designed highpass filter following the logarithm and an exponentiation stage. The frequency response of the filter should be circularly symmetric and real with a size invariance property determining its variation with distance from the origin. Objects * that are identical except for a change in size have Fourier transforms related by:

$$f_1(x,y) = f_2(ax, ay)$$

$$f_1(x,y) \overset{FT}{\longleftrightarrow} F_1(u,v) = F_2(u/a, v/a) \overset{FT}{\longleftrightarrow} f_2(ax,ay)$$

$$F_1(R,\theta) = F_2(R/a, \theta) \quad ; \quad R^2 = u^2+v^2, \quad \theta = \tan^{-1} v/u$$

And suppose the frequency response of the filter has the form:

$$H(R,\theta) = bR^d + c \quad ; \quad 1 < d < 0$$

---

* By the term objects we mean the logarithm of light intensity reflected from the actual physical objects.

Then objects after enhancement will have Fourier transforms given by:

$$f_1'(x,y) \xrightarrow{FT} F_1(R,\theta) \cdot (b \cdot R^d + c)$$

$$f_2'(\alpha,\beta) \xrightarrow{FT} F_2(\rho,\theta) \cdot a^{-d}(b \cdot \rho^d) + c$$

Note that the changes to F_1 and F_2 brought about about by the enhancement process are very similar. The only significant difference is a factor of a^{-d} indicating that the smaller object has all its Fourier components amplified more than the larger object by this factor. The constants b and c determine the high and low frequency gain and d is an empirically determined parameter which controls the flare effect which will be mentioned in connection with figure 11. The following values were used for the enhancement of figure 11:

$$b = 2 \cdot R^{-d} - c \quad ; \quad c = 0.5 \quad ; \quad d = 0.25$$

Reproducing enhanced images on photographic paper requires an appropriate scaling operation so that the usable light to dark range * of the paper is utilized efficiently. This may be done automatically if the image is first scaled by a constant, enhanced as above and then scaled by the reciprocal of the constant. If the

---

* In spite of the fact that typical photographic paper of the kind used in this report has a ratio of maximum to minimum reflectance of about 200:1, images are usually printed as light as possible without losing highlights.

constant is chosen properly, this step causes wide dynamic range (large ratio of light to dark in light intensity) images to be printed as light as possible while images with moderate to low dynamic range, which presumably are already quite light, have their average gray level modified to a lesser extent. The constant used for scaling in this experiment was Max(I)/4. Max(I) is the maximum intensity of the unprocessed image. Omitting this scaling operation causes the average gray level to be changed by the linear filter in an uncontrolled way.

Figure 10 is an outdoor scene with considerable variation in illumination. Objects outside the shed are lighted by direct sunlight and lighting in the shadow is by less intense scattered light. The same scene is shown in figure 11 after having been processed by an enhancement system based on the multiplicative model of figure 9.

If the model accurately represents processing done by the human visual system, then figure 10 can be considered to have been been processed only by the visual system while figure 11 has been processed externally by the enhancement mechanism as well as by the visual system. This comparison gives a qualitative indication of how the visual system is able to reject illumination effects. Note how the ratio of light in the shadow and outside the shed has been reduced and how small detail such as that on the boards has been made more prominent.

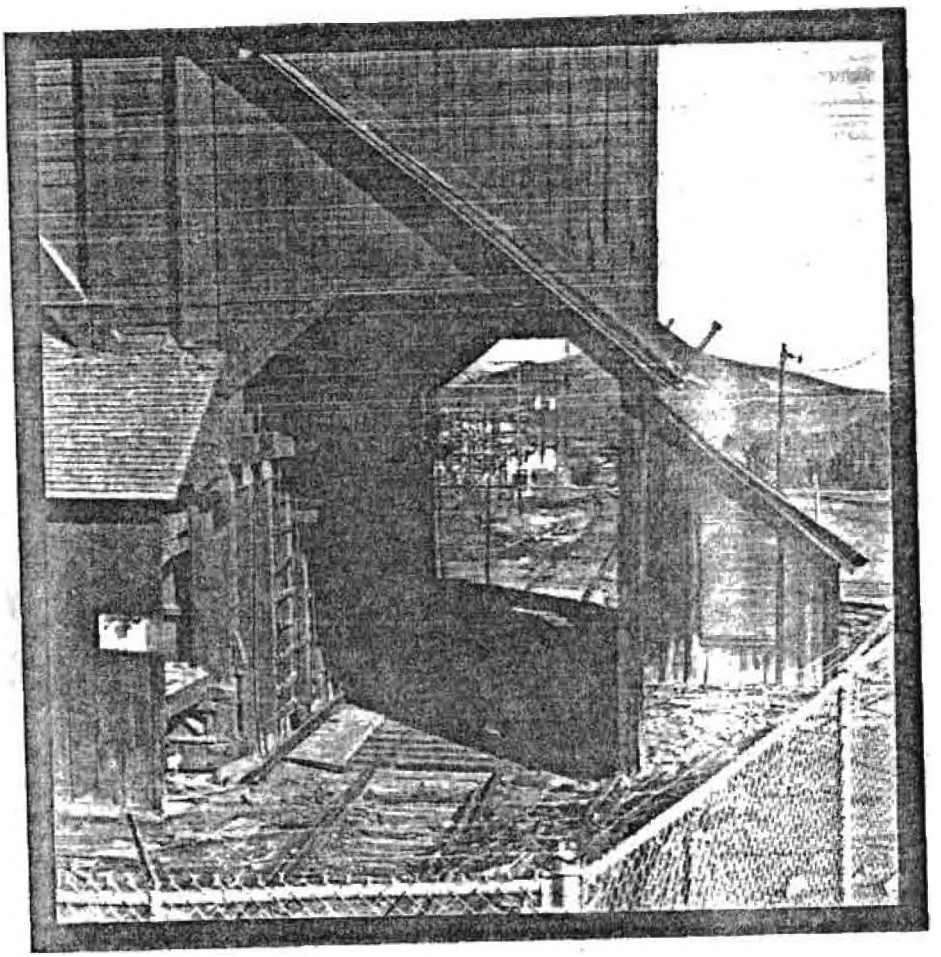


Figure 10 -- Wide dynamic range scene

The reader may notice a flare-like condition around the edge of the shadow in figure 11. A simple phychophysical test * will show that the effect is not generated by the visual system but is an artifact produced by violating the illumination smoothness condition

---

* Mach bands (figure 1) increase in strength as a the intensity gradient is made narrower and steeper, but they disappear when the gradient becomes a step. See Baudelaire [11] page 104 for a pattern which demonstrates this.

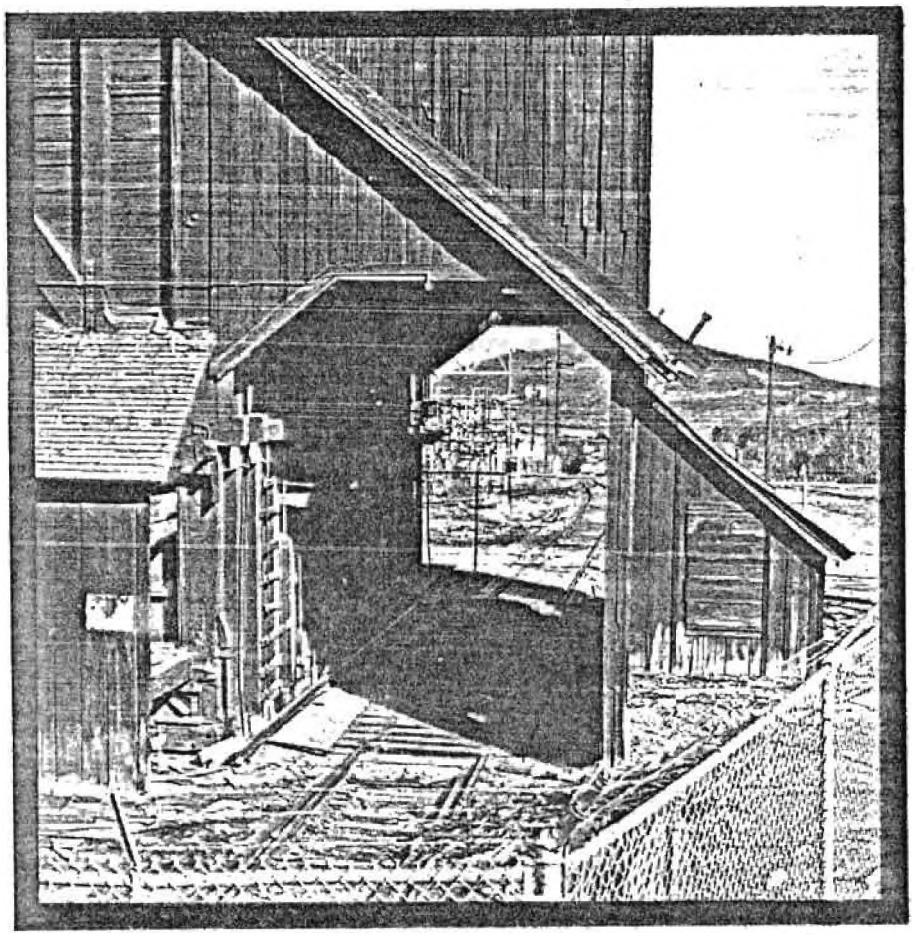


Figure 11 -- Multiplicative enhancement

and is nothing more than the step response of the linear filter. Apparently the visual system is capable of removing most of the illumination component from an image even if the variation in illumination varies step fashion as in figure 10. Since the visual system does not exhibit the flare effect, it is common [Baudelaire chapt. 5,11] to refer to the neural interaction in the retina as linear except near edges. In spite of this artifact, the enhancement of figure 11 is quite pleasing from an artistic point of view and

definitely gives the impression that reflectance information has been made more prominent while illumination effects have been suppressed. Other image processing operations based on the multiplicative model have been implemented successfully, notably coding and bandwidth compression [20,21].

The work described in the next chapter deals with other visual effects not accounted for by the multiplicative model. This work resulted in improvements which tend to moderate the edge artifact.

2.6 Illusions

The demonstration of figure 1 was cited by Mach as evidence of a second difference operator somewhere in the visual pathway. The model of figure 9 is similar to Mach's description and produces the bands he noted provided the linear filter has a reasonably smooth highpass frequency response and is approximately circularly symmetric. It is true that the model does not perform correctly near edges but it is capable of predicting successfully the illusions of figures 2 and 3. The remaining demonstrations require more sophisticated models of the type described in chapters 3 and 4 and are not predicted successfully by the multiplicative model.

2.7 Summary

A model of brightness perception consisting of a logarithmic stage followed by a highpass filter fits well into the context of Weber's Law and also predicts certain brightness illusions correctly such as Mach bands and the simultaneous contrast illusion. The

multiplicative property allows illumination and reflectance components to be processed separately suggesting a kind of image enhancement.

In the next chapter further physiological findings will be used as a guide in extending the model to deal with adaptation effects.

CHAPTER 3

A FREQUENCY SELECTIVE MODEL

3.1 Introduction

The results of recent contrast threshold experiments [12,21] and experiments involving electrical stimulation [23,24] of the visual cortex suggest that a process resembling a two dimensional Fourier analysis is performed by the visual system. This frequency analysis is helpful in explaining a wide class of visual phenomena including contrast threshold shifts and angular illusions. The model described in the previous chapter may be extended to account for these phenomena by adding a bank of bandpass filters connected to individual automatic gain control (AGC) elements * as shown in figure 12. This new model is a major contribution of the work described here. For images containing approximately uniformly distributed texture, the AGC elements pass the signal without change and the new model reduces ** to the multiplicative model of chapter 2, thus partially maintaining the multiplicative properties described there.

---

* The AGC elements are homomorphic systems of the type described by Stockham [24].

** A parallel combination of linear systems is linear with frequency response equal to the sum of the frequency responses of the individual systems forming the combination. The individual systems in this case are elements of a set of bandpass filters and their parallel combination is equivalent to the identity system.

It is not clear how information in the channel signals is actually recombined, so except in the case of angular illusions where a priority encoding scheme (see section 3.7) might be used, we will assume the signals are simply added to form a composite signal which is passed on to higher perceptual centers in the brain. In actuality both mechanisms may be present with only one in active use at a time.

Since our interest centers on implications of the model, discussion of electrophysiological evidence will be brief and most of the chapter will deal with two major contributions of this research. First, it will be shown how image enhancement based on this model can soften regions of strong texture while accentuating weak texture. The second contribution is a new method for removing unknown photographic blur which gives results superior to those previously obtainable. An interesting extra feature of the model is a plausible explanation for a class of angular illusions whose cause has been elusive. Arguments in favor of the filter bank * model are based on evidence from electrical probes of the visual cortex [23] which show that certain neurons are sensitive to stimuli based on their spatial frequency content. Adaptation experiments, some of which are demonstrated in chapter 1, also support the filter bank idea.

3.2 Evidence of Fourier Analysis

In the experiments of Hubel and Weisel [22] and Campbell [23],

---

* The filter passbands may be concentric rings of different radius or fan-shaped wedges of different orientations or a combinations of both types.

small electrical probes were inserted into individual cells of the

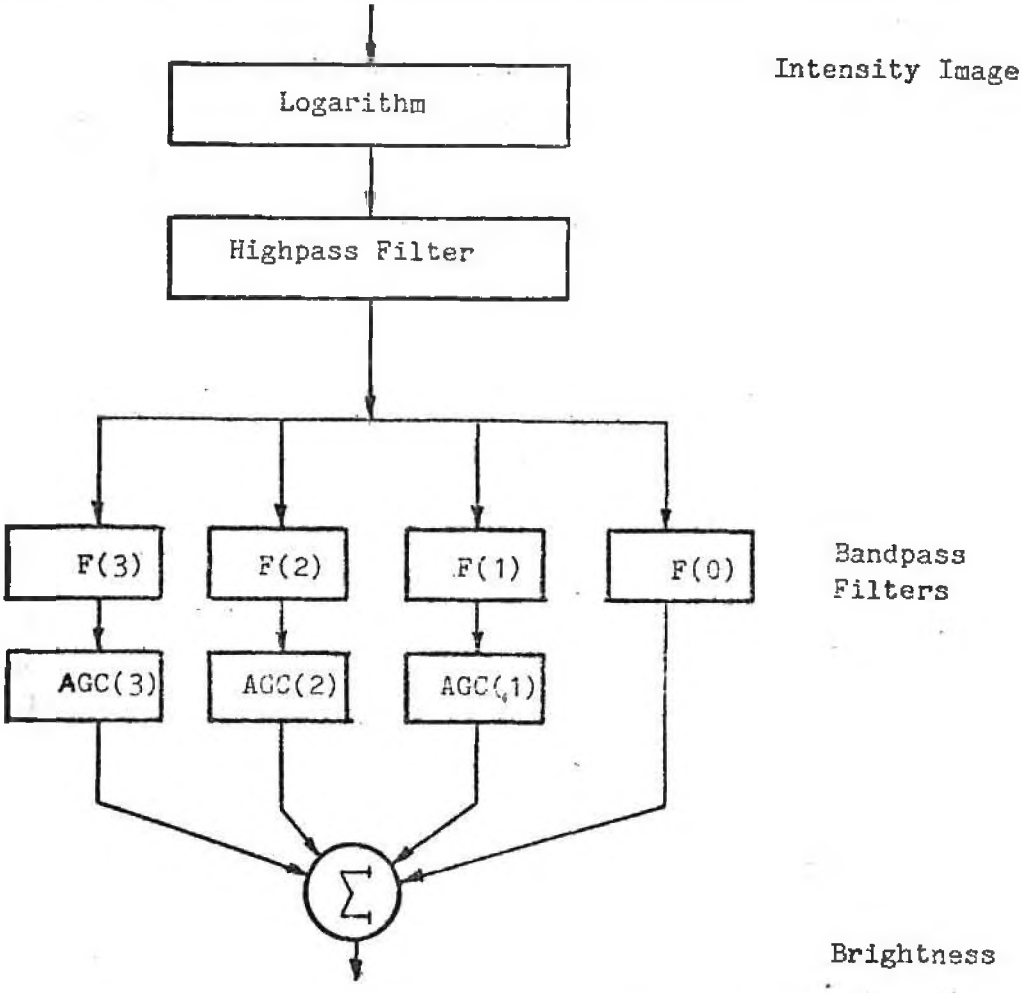


Figure 12 -- The frequency selective model of vision separates an image into channels based on spatial frequency content and does independent processing on each channel.

visual cortex in anesthetized animals to monitor neural response to various visual stimuli. In both studies a large number of cells were found to be sensitive to periodic bar gratings of a specific spacing and orientation. Often groups of neurons were found whose preferred angular orientation changed in a systematic way as the probe was moved from one neuron to the next. Most of these frequency selective

cells responded to bar patterns whose spatial frequency was within about an octave of the preferred frequency and whose angular orientation was within ± 15 degrees of the preferred direction. Some cells responded only to moving patterns thus indicating sensitivity to temporal effects. These, however, are outside the scope of the present study as are chromatic effects *. In most neural studies of this type there is an assumed similarity between neural organization in the animals tested, in this case monkeys and cats, and the neural structure in humans.

Psychophysical methods of probing the human visual system often rely on adaptation effects to identify particular neural structures. This is done by strongly stimulating the structure and noting any decrease in sensitivity. The motivation for this approach was described by Julesz [25] as follows: "if it can be adapted it must be there." Under this assumption, the presence of fatigue effects in what appear to be frequency sensitive neurons is an argument in favor of their existence. Another argument in their favor is the fact that a frequency selective model predicts results which closely parallel certain known aspects of human vision.

In an adaptation study by Blakemore and Campbell [19], prolonged viewing (45-60 seconds) of a high contrast sinusoidal grating was found to decrease the sensitivity to gratings of similar spacing and orientation by a factor of five. Gratings of a different spacing or

---

* See Frei [26]. At the time of this writing there is an active research program at the University of Utah dealing with color vision.

orientation did not produce the threshold shift and adaptation was not effective in changing the sensitivity to gratings whose spatial frequency was less than 1-2 cycles per degree, regardless of the spacing or orientation. Their experiments indicate that the spatial filters are quite broadly tuned, being about an octave in bandwidth with the lowest frequency filter being centered at 3 cycles per degree. All of the filters seemed to exhibit the same ±15 degree angular selectivity. These values of bandwidth and angular selectivity were used in the design of the image enhancement experiment described in section 3.4.

The gratings in Figure 4 may be used to illustrate this adaptation effect. Viewing the high contrast bars for 45-60 seconds causes a reduction in sensitivity to images containing similar Fourier components. The adapted frequencies correspond to every third, odd harmonic (third, ninth, ...) of the large, low contrast bars of figure (4b), and reduced sensitivity to these components causes the bars to appear blurred. For some observers the low contrast bars disappear completely for a short time after which they appear blurred and finally they return to their normal appearance.

The demonstration of figure 5 is not nearly as simple as the adaptation experiment noted above and an argument can be made that it is really nothing more than a laboratory curiosity, nevertheless, it is interesting in the context of a visual model having filters which are angularly selective and capable of being adapted, particularly if there is inter-channel communication causing the alternation.

Evidence cited above suggests further investigation into the

properties of a frequency selective visual model.

3.3 Texture Equalization

Easily the most noticeable property of this model is its variable sensitivity to detail. In a strongly textured environment, such as the apple blossoms of figure 15, the sensitivity is reduced because large amplitude channel signals cause a reduction in gain of the AGC elements. In the absence of detail the sensitivity is increased. The adaptation effect noted earlier is an example of this property. An aviator scanning the sky on a clear day for distant aircraft is an example of a conscious effort to exercise the variable sensitivity property, the goal being increased sensitivity. Covering minor flaws in a plastered wall with textured paint is an example of the visual system's reduced sensitivity to strong texture. An animal in a highly textured wilderness environment is less noticeable than if the environment were texture free. These examples of strong texture causing a reduction in sensitiviy to detail are similar to masking phenomena in hearing.

Campbell's contrast threshold experiment indicated that adaptation effects seem to suggest octave wide frequency bands starting at about 2.0 cycles/degree. Using his experimental data as a guide, four overlapping filters were used in the enhancement experiment. A low frequency (0 - 2.0 cycles/degree) channel was used without AGC, and bandpass filters centered at 3.0, 6.0, and 12.0 cycles/degree were used with AGC as shown in figure 12. Since the Nyquist frequency for a 3" x 3" image viewed at arms length is

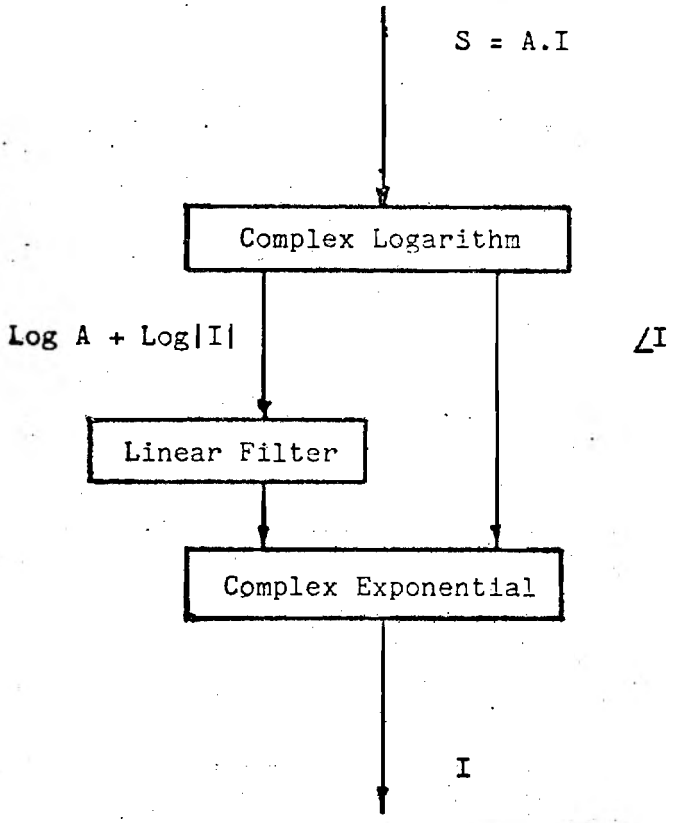


Figure 13 -- Homomorphic AGC element

approximately 15 cycles/degree when it is sampled on a 256 x 256 grid, four filters are all that are required. Each filter was circularly symmetric with zero phase.

Let us supose that signals emanating from the bandpass filters consist of two components combined by multiplication; one of them, which we will call the amplitude signal, is always positive, and the other signal may be both positive and negative, zero mean and may be thought of as the information signal. In figure 15 the amplitude

signal is small in the vicinity of the clouds and large in the apple blossoms. Whenever texture is strong the amplitude signal is large and when the texture is weak the amplitude signal is small. Taking the complex logarithm converts this product into a sum whose constituents may be separated if their spectra are disjoint.

There are two important parameters associated with the linear filter of figure 13. Attenuation of the amplitude signal depends on the low frequency gain which may be thought of as controlling the "aggressiveness" of the AGC unit. Figure 17 was processed with the low frequency gain set at 0.25. The cutoff frequency is set to a value slightly less than twice the lowest frequency passed by the bandpass filter. Note that the gain for zero frequency (R = 0) is equal to unity. This has the effect of preserving the average texture strength across the entire image.

It should be clear that attenuation of the amplitude signal by a properly designed AGC unit will equalize texture across an image.

3.4 Image Enhancement

An image enhancement experiment based on the texture equalization property shows how regions of strong and weak texture are modified by the model. Figure 15 is a natural scene containing regions of strong texture (the apple blossoms) and regions of weak texture (the clouds). The next figure is a multiplicative enhancement of the type described in chapter 2. Detail in the clouds is still relatively weak while the apple blossoms are stronger than in figure 15. Note that this scheme makes adjustments in contrast

based primarily on the basis of texture size. Figure 17 is a texture equalization type enhancement.

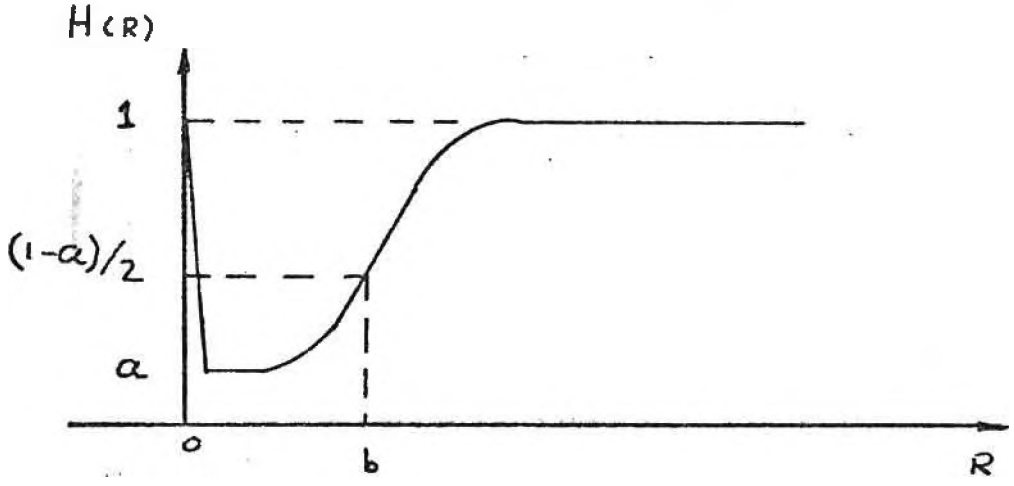


Figure 14 -- Radial frequency response of the AGC filters

Information conveyed by weak texture such as the indentations in the dome and detail in the clouds is much more visible than in either of the two previous figures. Note also the softer treatment of the apple blossoms. It might be said that figure 17 is more "readable" than figure 15 or 16 in the sense that subjectively important information is more prominent. Adjustments in contrast are now

dependent on strength of texture as well as on size. It is, of course, possible to find fault with figure 17 from an artistic point of view since the AGC units used in the experiment were quite aggressive (see figure 14). The purpose of this experiment was to demonstrate the texture equalization property and one might expect a less aggressive system to give more pleasing results.

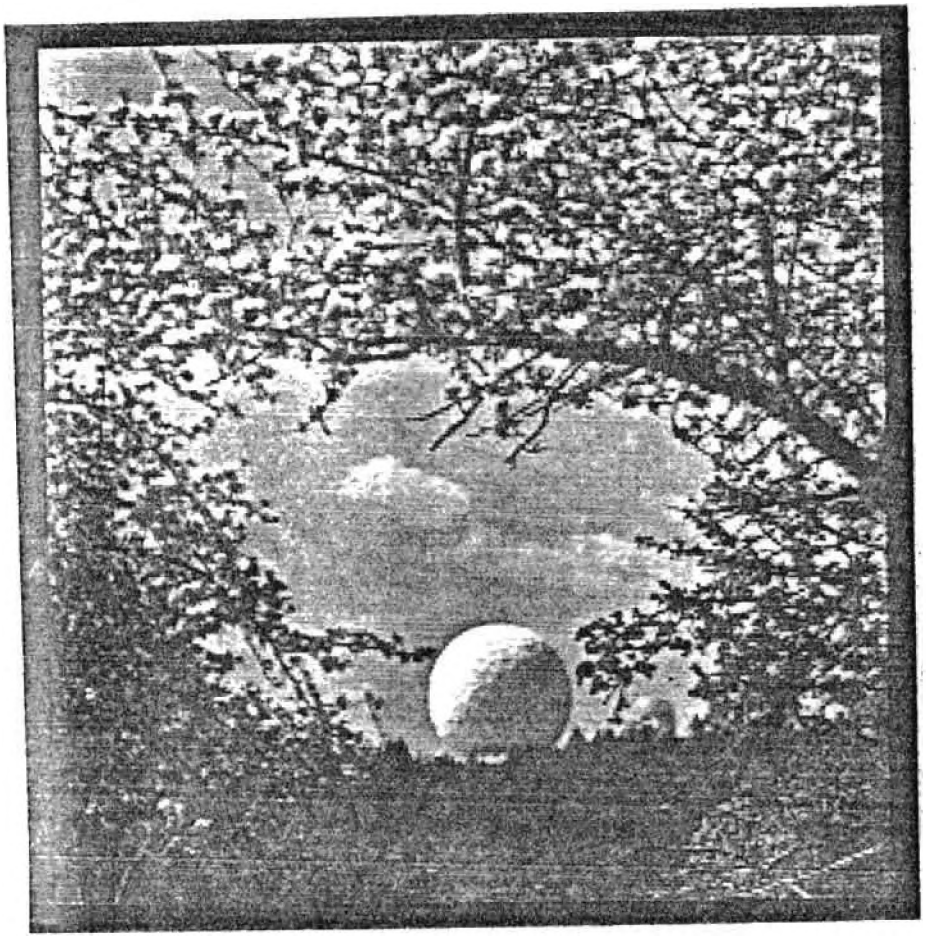


Figure 15 -- Natural scene containing variable strength texture

The texture equalization property is a result of frequency

selective channels in the visual system acting in concert with an adaptation mechanism. Electrophysiological experiments cited earlier assoiate this process with frequency selective neurons in the visual cortex, and adaptation experiments have provided us with a quantitative description of filters which simulate the effect. Fatigue effects such as numbness and auditory threshold shift

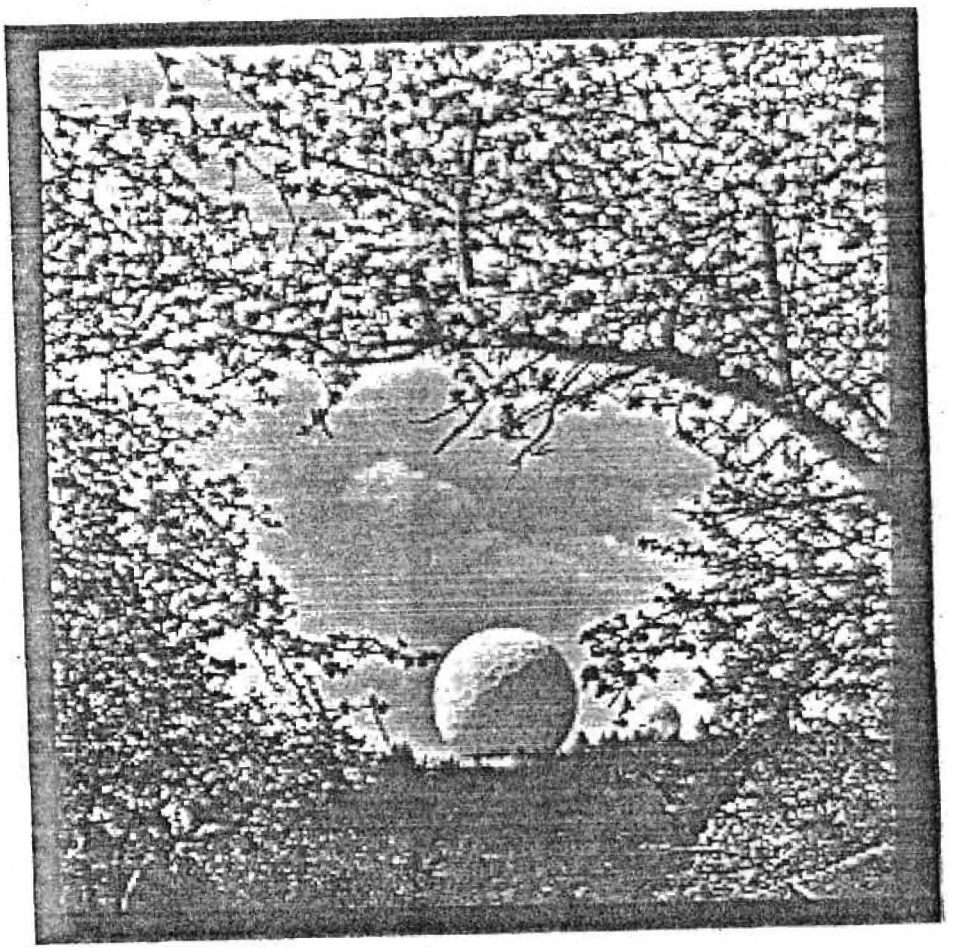


Figure 16 -- Enhancement of figure 15 based on the multiplicative model. The apple blossoms, which were quite prominent in the original (figure 15), are further accentuated by the enhancement operation.

are common in sensory perception, and the adaptation to periodic bar gratings by the visual system may depend on the same kind of neural mechanism. Incorporating these ideas into an image enhancement mechanism has shown how weak detail may be accentuated while keeping strong detail from being amplified excessively.

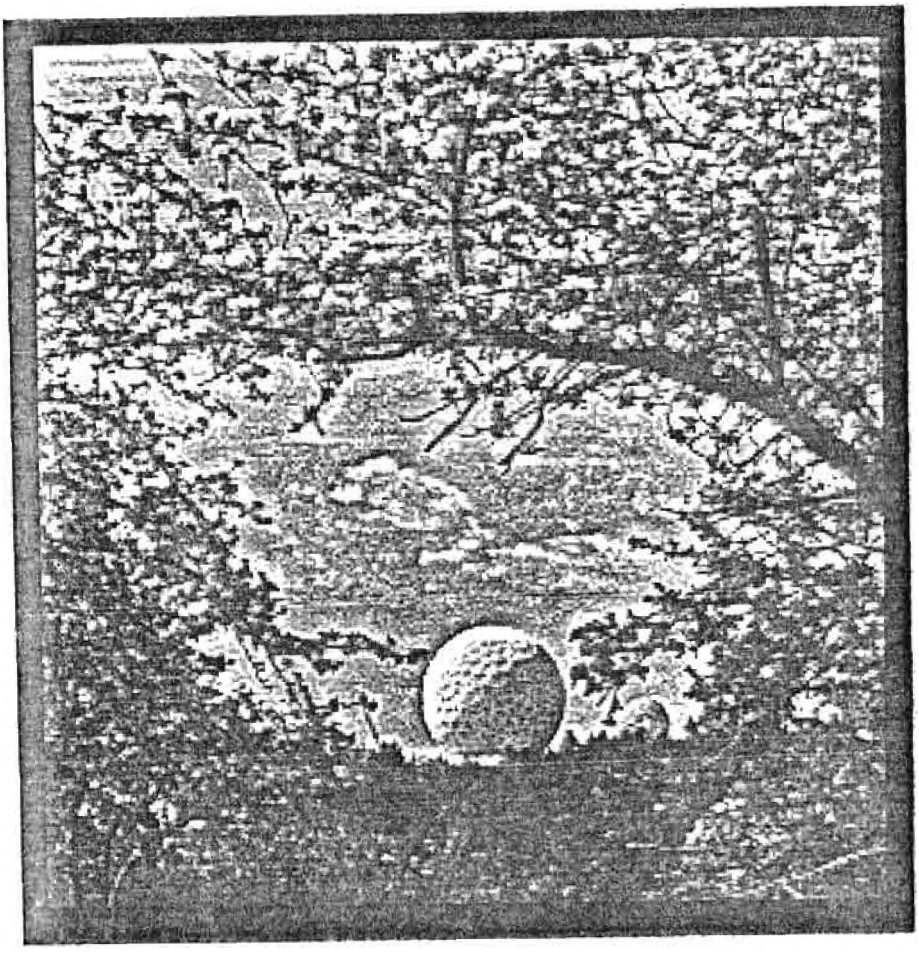


Figure 17 -- Image enhancement based on the frequency selective model results in a kind of texture equalization where weak detail in the dome and sky are strongly accentuated relative to corresponding parts of figure 15 while the apple blossoms are given a softer rendition.

The edge artifact noted in chapter 2 is moderated by texture equalization since the channel signals, which will be large near edges, will be attenuated by the AGC elements.

In the next section, we will describe a method, based on the texture equalization property, for restoring blurred images.

3.5 Restoration of Blurred Images

A little reflection should suggest a similarity between blurred photographs and images containing weak high frequency texture, and one might expect a system capable of accentuating weak texture to sharpen a blurred photograph.

The ideas upon which this section is based first came to mind by observing that each AGC element in the frequency selective model has associated with it a criterion for determining when the signal in its channel is too weak and must be amplified. This means that the local power spectrum * is being adjusted to some previously specified prototypical value. These ideas led to the deblurring mechanism in figure 18.

The method will be described as concisely as possible despite the fact that in certain respects its mathematical foundations are rather subtle. Following this discussion, some comparisons between the deblurring method and the frequency selective model will be made which suggest a limited deblurring capability in human vision.

---

* The idea of a local power spectrum is similar to the short time spectrum on which speech spectrograms are based.

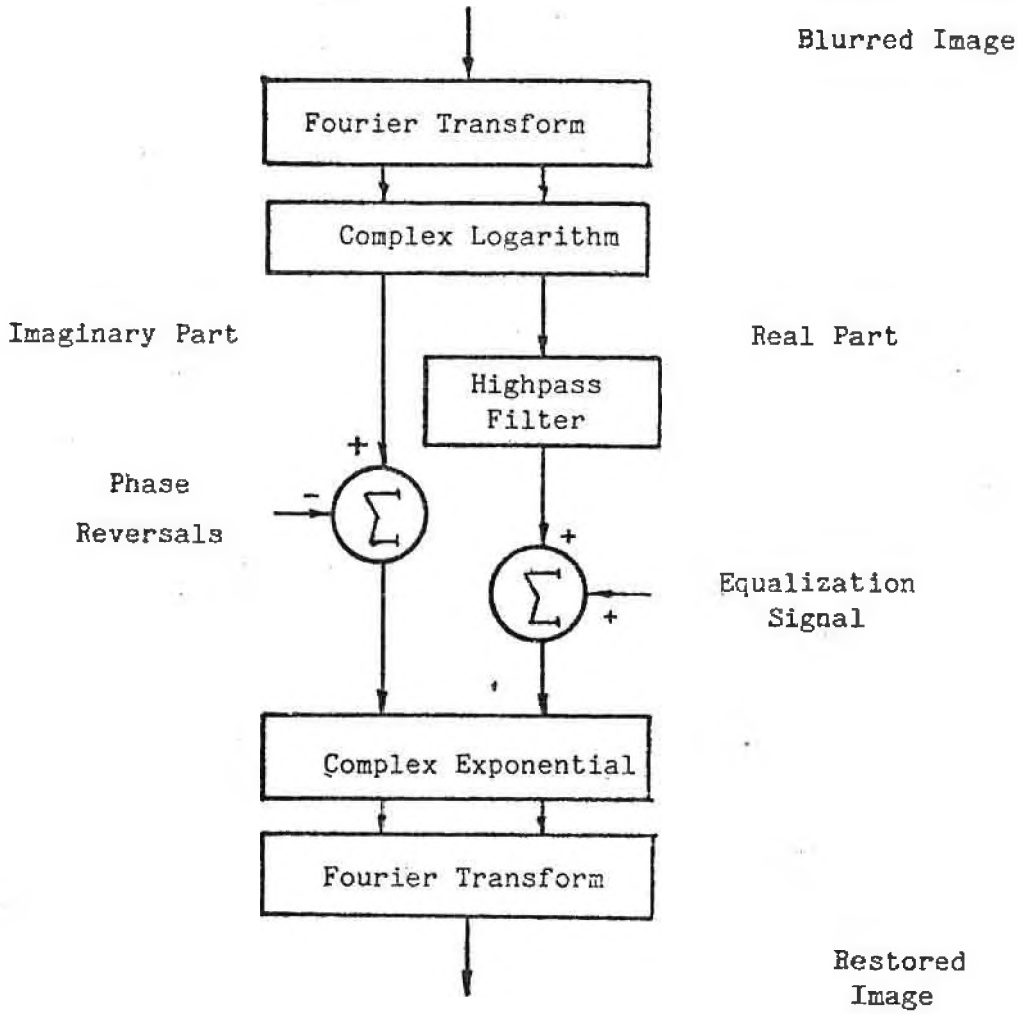


Figure 18 -- A deblurring mechanism suggested by the texture equalization property.

Blurring caused by camera motion or an out of focus lens are processes that combine a clear image with a blur impulse response by convolution. An out of focus blur has a cylindrical impulse response and a motion blur has a fence-like impulse response. Stylized examples are shown in figures 34 and 37. The process of deconvolution [2] is the process of removing one of the components

(the blur) by first mapping the blurred image into a space in which the two components are added rather than convolved, and then the blur information is removed by linear filtering *.

Several obstacles stand in the way of what might seem otherwise to be a straightforward task:

(1) The first complication involves noise-like fluctuations, known as film grain, introduced when images are recorded on photographic film. If the restoration is to be acceptable these random fluctuations must not be amplified excessively by the restoration process.

(2) The second difficulty arises because space-limited blurs of the type illustrated above completely eliminate image energy at certain spatial frequencies making an exact restoration impossible even in the absence of film grain noise, however if there is no noise and only a few frequencies are eliminated, their absence goes almost unnoticed. Figures 19-23 illustrate this.

(3) Approximate aperiodic convolutional inverses exist for space-limited blurs but are often nonzero ** over large domains ***.

---

* There is a terminology associated with this style of signal processing in which the filtering just described is known as liftering, and it is implemented by multiplication in the so called cepstral domain. The independent variable in this domain is called quefrency and is described as being short or long rather than high and low as is the case for the frequency domain. Using these terms figure 18 is really a long pass lifter. For the purposes of this report the more familiar terms Fourier transform, logarithm and filter will be used to describe these processes.

** See the footnote in Appendix B for a interpretation of the term nonzero.

*** See Appendix A for an illustration of this difficulty.

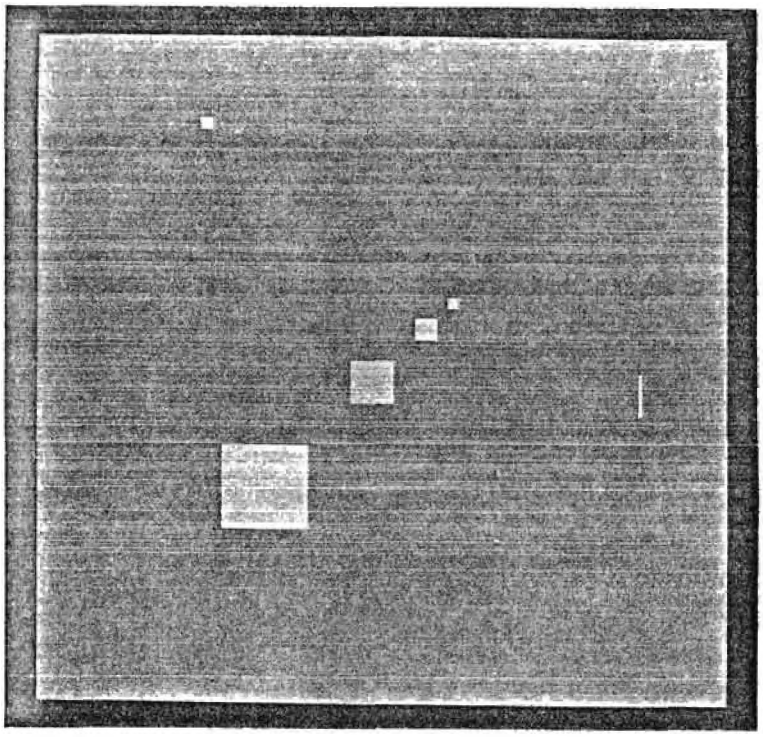


Figure 19 -- Artificially created test image

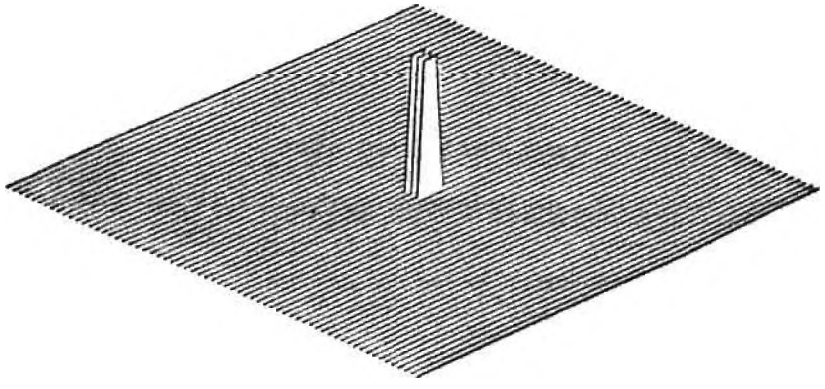


Figure 20 -- Artificial blur impulse response

(4) The fourth problem is caused because the blurred image is truncated abruptly at the edges of the film by the camera's film holder. Special edge treatments are required to suppress artifacts generated by these edges.

Item mumber three can be particularly troublesome because restoration filters for use with aperiodic convolution often have very long impulse responses, and arbitrarily truncating them for computational convenience causes severe ghost-like echoing in the restored image. Figures 24 and 25 illustrate this problem for the artificial blur. Traditional attempts to remedy the situation by windowing have proved to be only partially successful. See Cannon [4] page 34 and Cole [5] page 53 for examples of echo-like artifacts introduced in blur removal schemes by impulse response truncation.

An alternative to windowing is to take advantage of the fact that both the blurred input image and restored image are usually of finite size (often the same size) and a convolutional inverse can be truncated without introducing echoes provided that its domain is adequate *.

The approach to the truncation problem taken here relies on the limited spatial extent of most blur impulse responses which allows substitution of periodic ** for aperiodic convolution.

---

* See Appendix B for a discussion of constraints on the impulse response length imposed by this technique. This approach may be impratical because of the large computational effort involved in obtaining the convolutional inverse prior to truncation.

** Appendix C shows how the effects of circular convolution may be obtained by filtering the log spectrum.

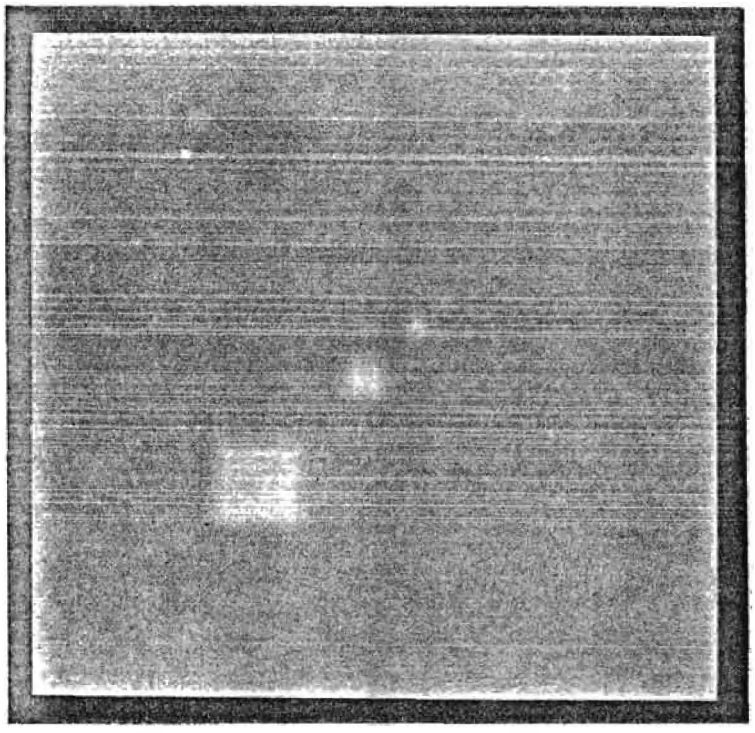


Figure 21 -- Test image artificially blurred using circular convolution (noise free)

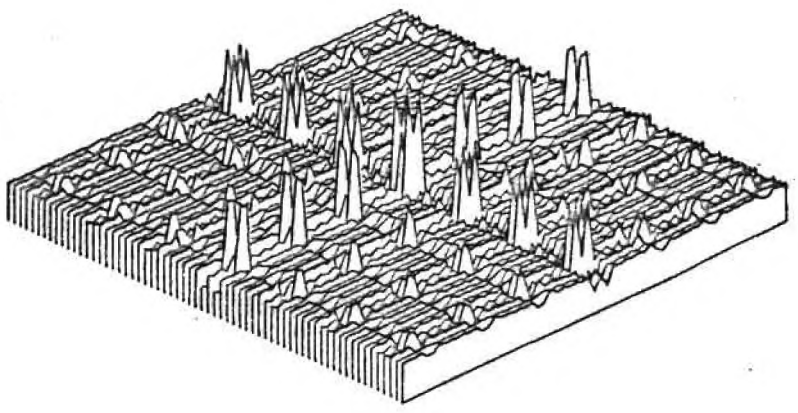


Figure 22 -- Convolutional inverse modified to avoid infinite gain at spatial frequencies where the blur system has zero gain.

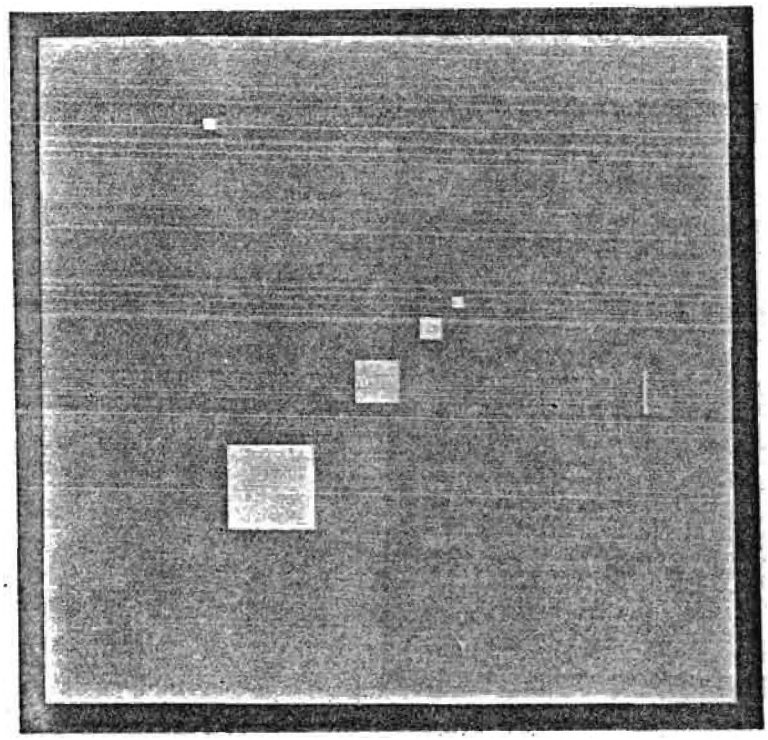


Figure 23 - Artificial image restored by inverse filtering

The blur systems considered here completely eliminate signal energy at certain spatial frequencies and severly attenuate it at others nearby. In attempting to amplify these frequencies adequately the system is constrained by the highpass filter so that noise is not amplified excessively. The next four figures should help illustrate this.

The success in substituting periodic for aperiodic convolution is because both kinds give identical results, except possibly near image boundaries. Note that results identical with figure 21 are obtained if figure 19 is extended with its boundary value, combined with figure 20 by aperiodic convolution and the result truncated to the original size.

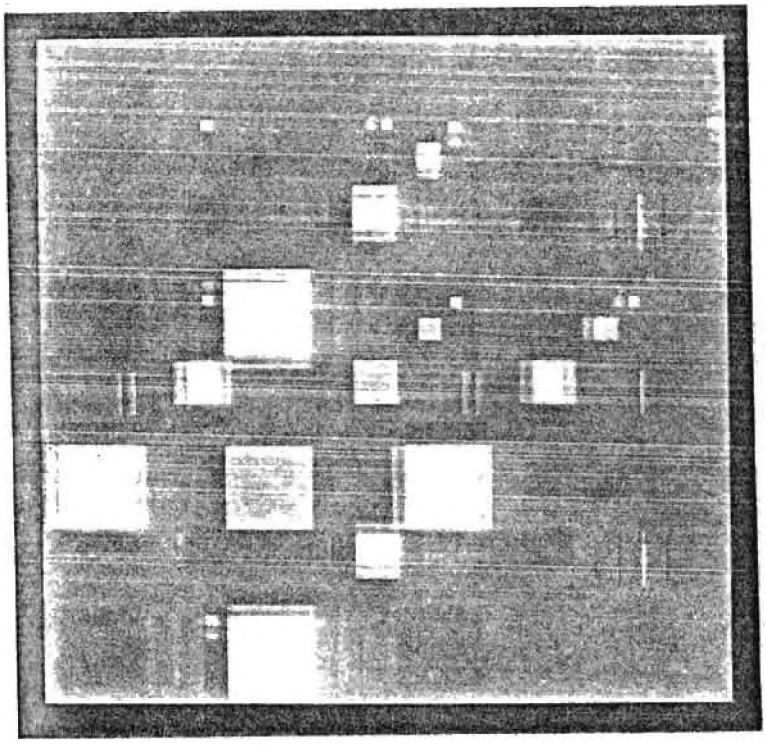


Figure 24 -- Artificial image restored using a convolutional inverse truncated to half its original size. Truncation caused large errors in the average value requiring manual corrections to make the image printable.

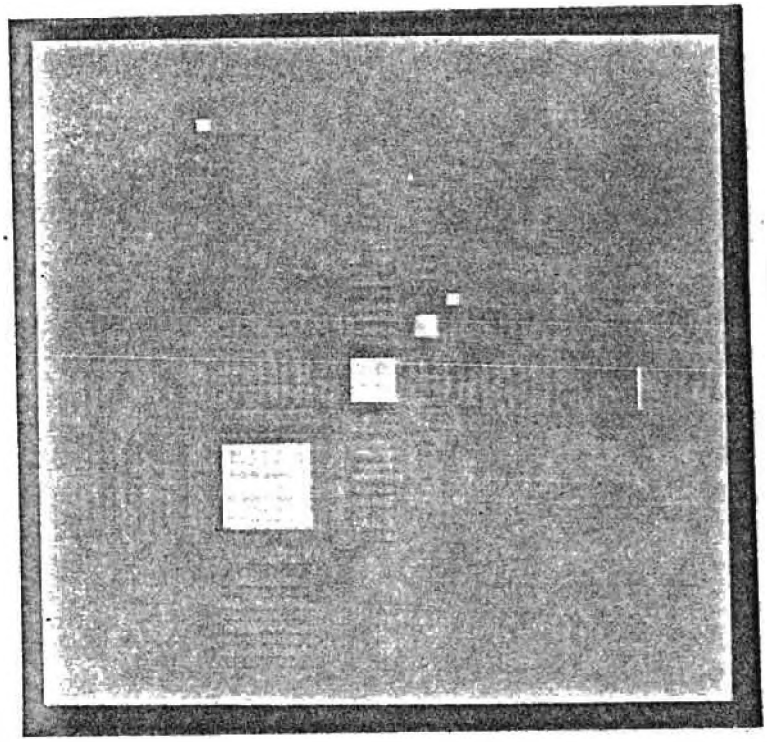


Figure 25 -- Same as figure 24 except the truncated inverse impulse response was windowed to reduce echoing.

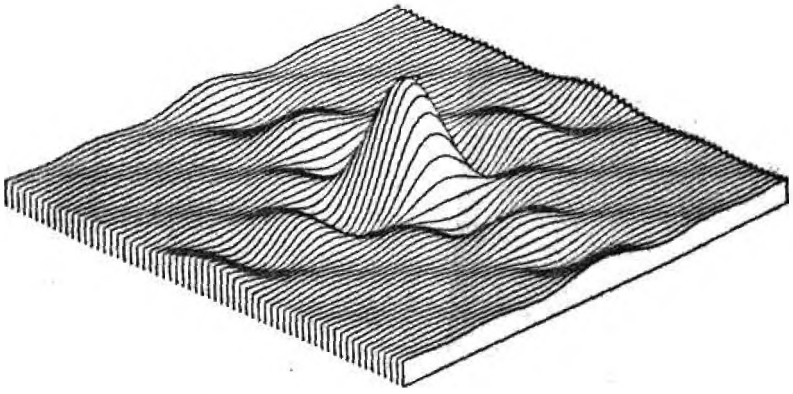


Figure 26 -- Blur frequency response

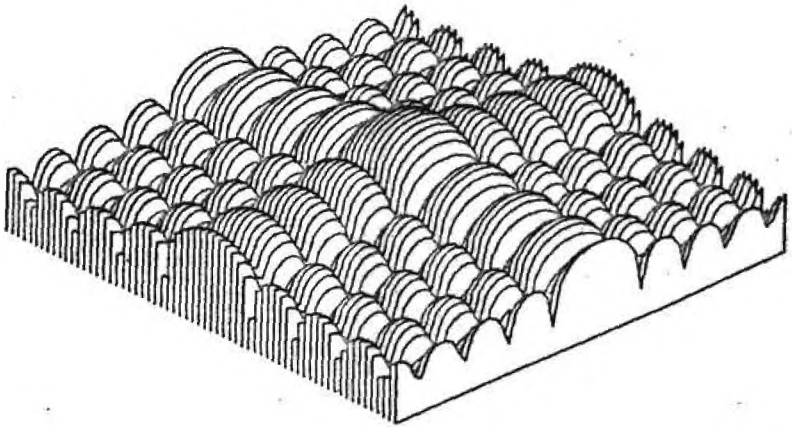


Figure 27 -- Log spectrum of the blur

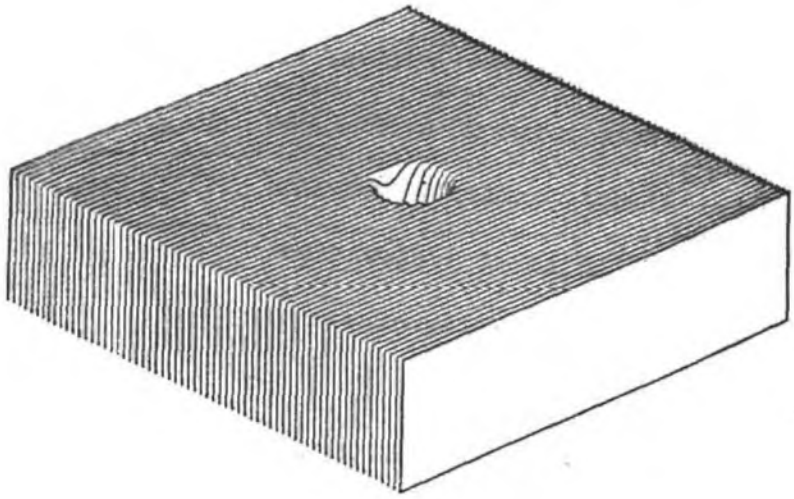


Figure 28 -- Highpass filter frequency response

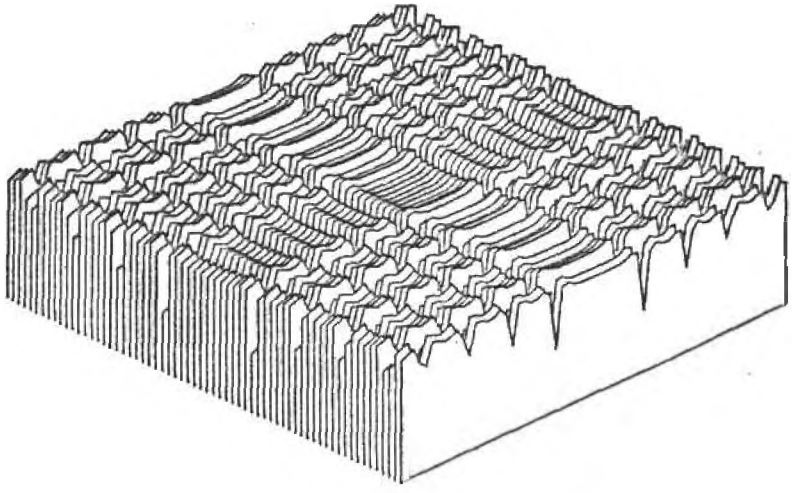


Figure 29 -- Filtered log spectrum

Figure 26 shows the frequency response associated with the artificial blur referred to earlier. Frequencies where this function is zero are the same ones where its logarithm (figure 27) increases negatively without bound. It is the task of the highpass filter to remove as much of this signal as possible while preserving regions of large negative value to prevent excessive amplification of noise. In figure 29 the blur log spectrum * has been almost entirely removed except in the region of potentially noise dominated frequencies. This has the effect of restoring the image while preventing excessive amplification of noise.

This same filtering effect is obtained if image information is present along with the blur, but part of the image is removed by the filter and must be reinserted in the form of an equalization signal like the one in figure 30. One which is properly prepared can be used in restoring a variety of blurred images. One way to prepare an equalization signal is to lowpass filter the log spectrum of a similar, unblurred image using a filter whose frequency response is related to that of the highpass filter as follows:

$$LPF(\omega) = 1(\omega) - HPF(\omega)$$

Filters of this type are called complementary filters.

Phase information of the blur system is not present in figure 27 or 29 and no method of removing it has been described so far. In the case of the artificial blur, the phase is known exactly and

---

* Log spectrum is used to mean the real part of the complex logarithm of the discrete Fourier transform.

subtracting it is straightforward. For real blur systems encountered in the field a method such as the one by Cannon [4] may be used to estimate the phase of the blur after which it may be subtracted. Figure 31 shows the phase associated with the artificial blur.

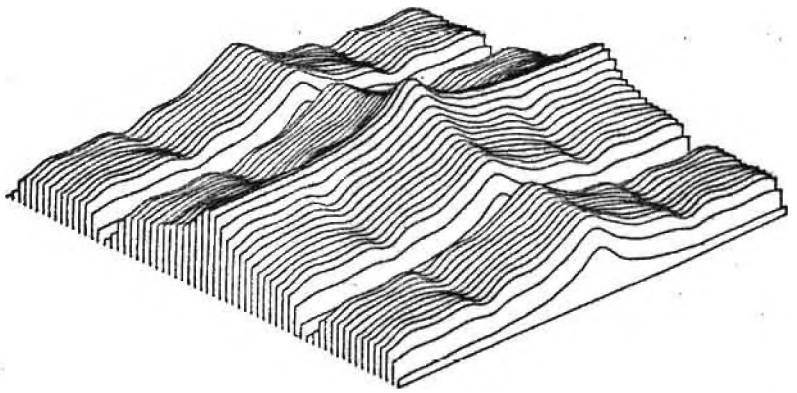


Figure 30 -- Equalization signal

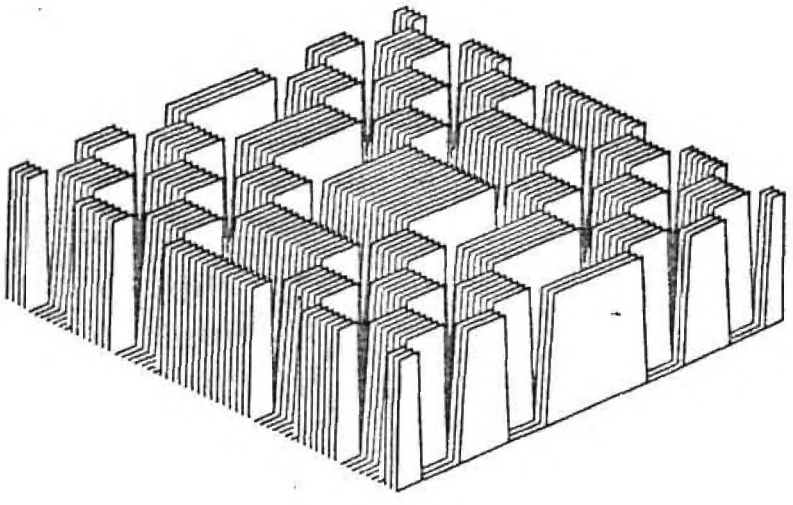


Figure 31 -- Phase of the artificial blur

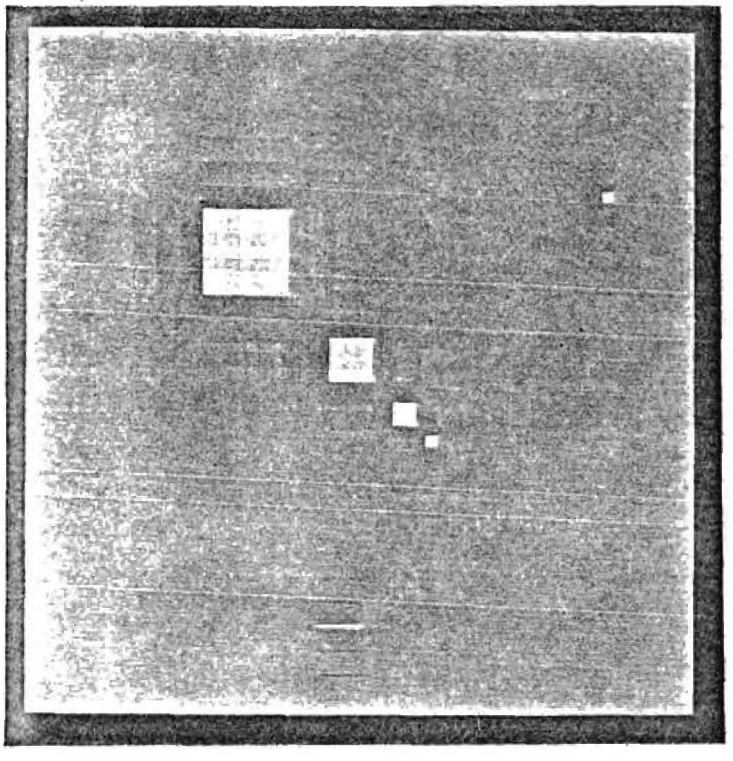


Figure 32 -- Image restored by the method of figure 18

In figure 32 the artificially blurred image is restored using the techniques described above. There are two reasons why the retoration is not identical with the one in figure 23. First, the effects of noise are anticipated but none is present. This capacity for dealing with nose is not needed but it constrains the restoration as if it were needed. Second, the exact details of the blur log spectrum were unknown to the restoration system. This is known as "blind deconvolution."

Figure 33 -- A roadside sign blured by an out of focus lens.

Applying these ideas to images blurred in the field requires a special edge treatment to make the image have the same value along its boundary as shown in figure 36. The blurred image is treated as though it has been replicated periodically and a special interpolation process * is invoked near image boundaries to simulate periodic blurring across the boundaries. In this setting, questions regarding convolutional inverses having large spatial extent do not arise since the inverse is periodic and must be determined only for a single period. If the edge treatment is successful no further attention is required with regard to problem four (truncation after blurring).

Figures 33-39 give an idea of the degree to which a successful restoration of images blurred in the field is possible using this method. Note how the small text in figure 36, which was blurred beyond legibility, has been made legible in the restored image of figure 38. Also the outline of the sign is much more clearly defined in both figures 35 and 38. The speckled appearance of figures 35 and 38 is due to amplification of film grain noise and may be traded for sharpness by adjusting the highpass filter cutoff frequency.

---

* The interpolation process consists of the following:

a. Removing the average value
b. Multiplying the image by unity everywhere except near edges and there by half of a Hanning window
c. Reinserting the average value.

This scheme simulates blurring across image boundaries.

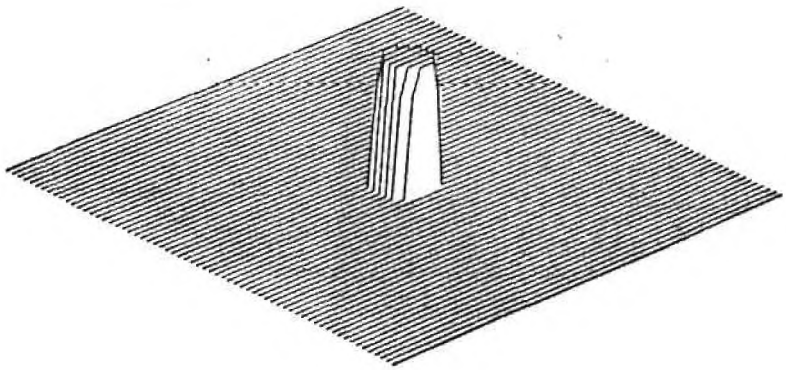


Figure 34 -- Stylized impulse response corresponding to the blur of figure 32

A minor artifact, barely perceptible in figure 38, is a shadow manifesting itself as a dark copy of the letter "c" to the left of the word convention. This may be due to regularly spaced harmonics that are missing from the restored image or perhaps errors in estimating the phase of the blur.

Figure 35 -- Restored image corresponding to figure 33

Figure 36 -- Sign blurred by camera motion

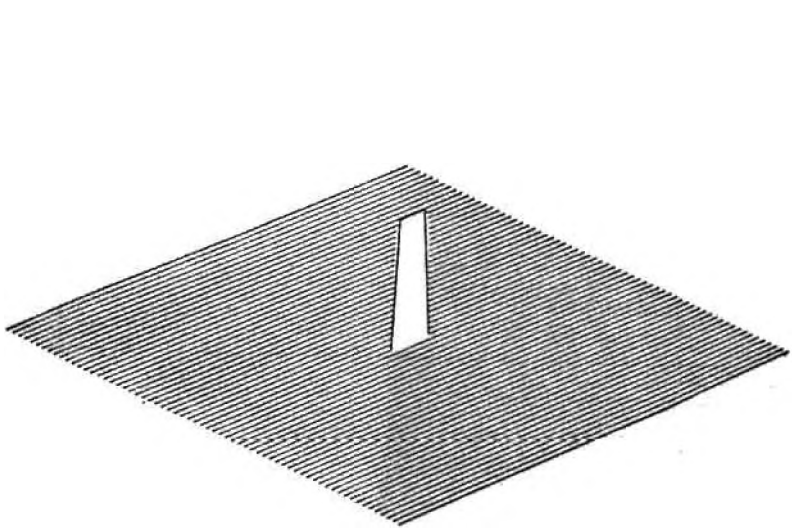


Figure 37 -- Motion blur impulse response This figure represents a function that is zero except along a line in the center where it is a constant.

Figure 38 -- Restored image by the method of figure 18. Note the effects on the right and the left edges of the images due to the interpolation process.

Figure 39 -- Sharp image of roadside sign

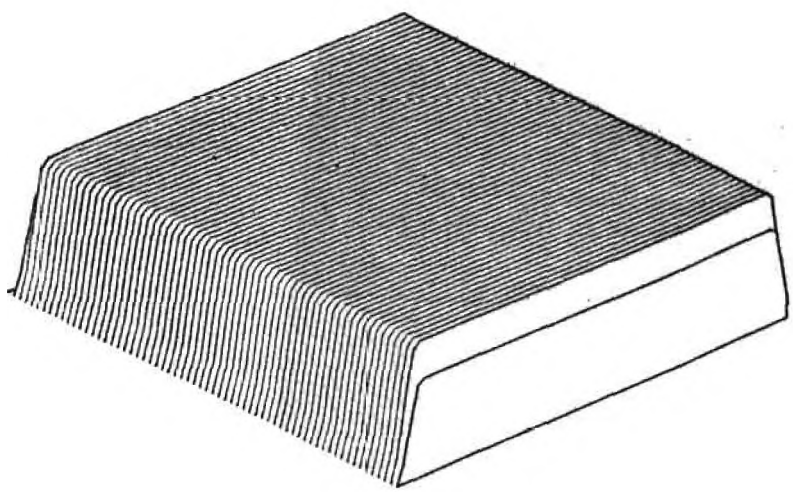


Figure 40 -- Interpolation function used to make the blurred image of figure 33 have a constant value at the boundary

Several similarities between the frequency selective model of vision (figure 12) and the deblurring mechanism (figure 18) suggest a limited deblurring capability in human vision. The Fourier transform is often naively understood in terms of a bank of bandpass filters similar to those in figure 12, a principal difference being the lower frequency resolution of the filters in the visual model. The AGC elements in figure 12 were implemented using Stockham's method [24] for separating multiplied signals and is similar to the logarithm, highpass filter, and exponential in the deblurring mechanism. One may view the equalization signal of figure 18 as providing a fixed gain for each frequency channel. There are only two places where the analogy breaks down. There is no logarithmic stage operating on the blurred image prior to the Fourier transform in figure 16, and the visual model has no provision for correcting phase reversals introduced by the blur. For images of low dynamic range such as the sign in figure 33, the logarithm is approximately linear and its absence is not critical. Failure to correct phase reversals is a significant difference between the two systems and is a major limitation on the ability of the visual model to restore blurred images.

The existence of this capability in human vision might help explain why it is difficult to make improvements in blurred photographs that are as striking as might be expected from physical considerations.

3.6 Angular Illusions

An interesting class of angular illusions is predicted correctly by a slightly modified version of the frequency selective model. The modifications are biologically plausible requiring only inhibitory relationships like those thought to be responsible for Mach bands, and a scheme for encoding angular information.

Suppose the vertical lines of figure 6 cause almost equal neural activity in three sets of channels whose preferred orientations are, slightly clockwise from vertical, vertical, and slightly counter clockwise. If these three channnels are inhibited by another channel which responds strongly to diagonal lines; and activity is most reduced in the channel whose orientation is nearest that of the diagonal pattern and least reduced in channels oriented away from it, then the illusion is predicted correctly. If it is also true that the preferred orientation of frequency selective neurons changes with their location in a geometrically sturctured way [22,23], and if angular judgements are made on the basis of which channel responds most strongly, the parallel lines will always appear rotated in the direction toward the perpendicular to the diagonal lines. This is exactly what happens in this class of illusions.

3.7 Summary

In this chapter experimental evidence favoring a frequency selective model of brightness perception has been described with emphasis on the model's texture equalization property. This property follows from electrophysiological as well as psychophysical

experiment and forms the basis of an enhancement process where changes in contrast are made on the basis of texture strength as well as texture size. A successful method for removing unknown photographic blurs based on this property was described that is very similar in its organization to that of the visual model. As a result of this similarity, a limited capacity for restoring blurred images was claimed for the visual system. Another satisfying property of the frequency selective model is its ability to predict correctly certain rather strong angular illusions.

In the next chapter a class of edge related illusions is considered in relation to a rather different model of brightness perception.

CHAPTER 4

THE MODIFIED RETINEX MODEL

4.1 Introduction

An alternative to the multiplicative model which explains the visual system's ability to reject illumination effects involves a kind of edge detector mechanism first proposed by Land [6,7]. Land's model correctly predicts the Cornsweet illusion but does not readily lend itself to image processing because it is not suitable for use on two dimensional sampled functions. A modification described in connection with the color constancy experiment overcomes this difficulty.

4.2 The Retinex

The term retinex was intended by Land to convey the impression of a neural network in the retina, optic nerve, and visual cortex for extracting lightness * information from images. To do this, ratios of light intensity are formed at adjacent points along a closed path and if a ratio is different from unity by only a small amount, it is set equal to unity. To extract lightness information from an image these

---

* Lightness is Land's term for the psychophysical correlate of reflectance and is similar to the term brightness as defined in chapter 3. However lightness is equivalent to surface reflectance and brightness is defined in terms of psychological responses to light stimuli.

ratios are multiplied together around the path. The region crossed by the path corresponding to the largest accumulated product is assumed to be white and is given an arbitrary value against which the lightness at other points is compared. Necessary and sufficient conditions for lightness information computed in this way to be independent of the path are, that the image consist of patches each having uniform reflectance, and that the path cross the same reference, white patch.

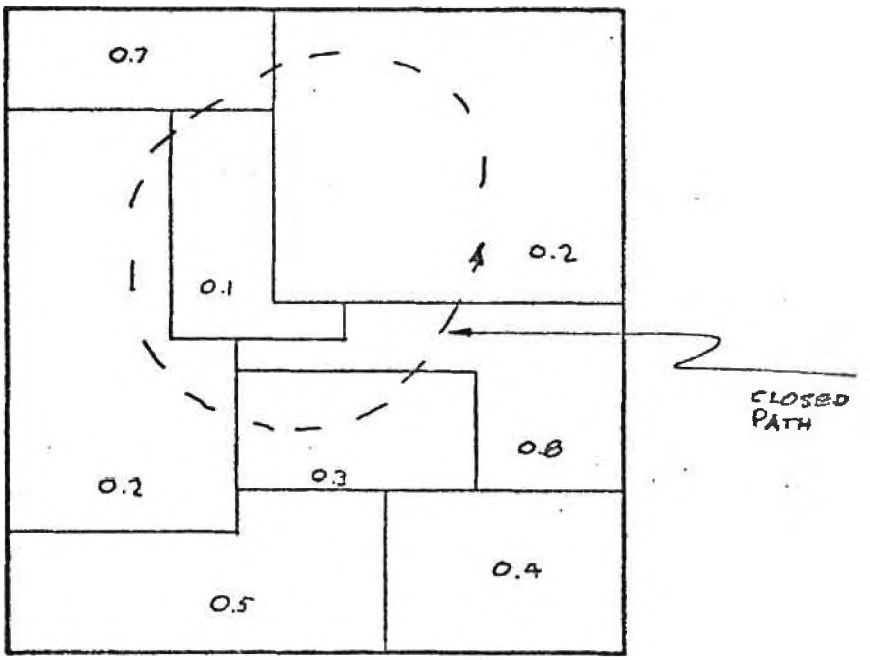


Figure 41 -- The retinex model as described by Land is capable of removing illumination gradients by a threshold process operating on ratios of reflected light.

Since in Land's formulation the path is arbitrary, computational complexity is introduced which makes the extraction of lightness quite difficult. A modification shown in figure 42 overcomes this problem.

It is easy to see how such a system might reject information about the average level of illumination since all lightness values are scaled in a way that assigns reference white a predetermined value. Lightness values thus depend on ratios of reflected light rather than on the illumination level.

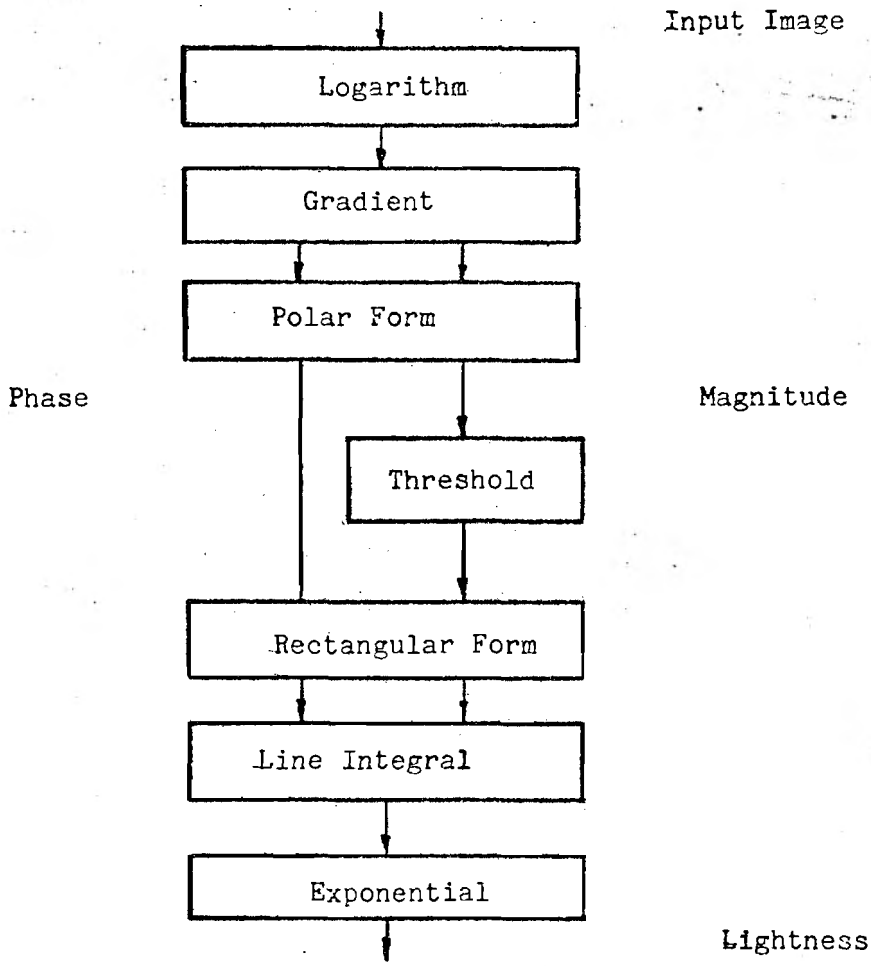


Figure 42 -- The retinex modified for image processing

Gradual changes in illumination across an image are also rejected by the retinex because light ratios within a patch will be close to unity and the threshold will set the ratio equal to unity. This eliminates any information about gradual illumination gradients. The color constancy experiment of section 4.3 illustrates both of these properties. The new implementation of Land's retinex consists of taking the logarithm and applying a threshold to the magnitude of the gradient rather than applying a threshold to ratios around a closed path. This permits computations to be carried out on a rectangular grid in a systematic way. Integrating the gradient may be done to within an additive constant by an appropriate * line integral, and the constant is supplied in the process of assigning a predetermined lightness value to reference white. A color constancy experiment will be described next which illustrates how the retinex can extract lightness information in the presence of illumination gradients.

4.3 A Color Constancy Experiment

Our interest in this experiment stems from the ability of visual system to perceive colored objects in correct color relationship to each other even under widely varying conditions of illumination. For example, a blue necktie looks nearly the same indoors under yellowish incadescent lighting as it does outdoors under much bluer natural light. This ability is known as color constancy.

---

* In the color constancy experiment an average of integrals taken over many paths of integration was used to minimize the effects of noise. See Wylie [27] for a discussion of line integration.

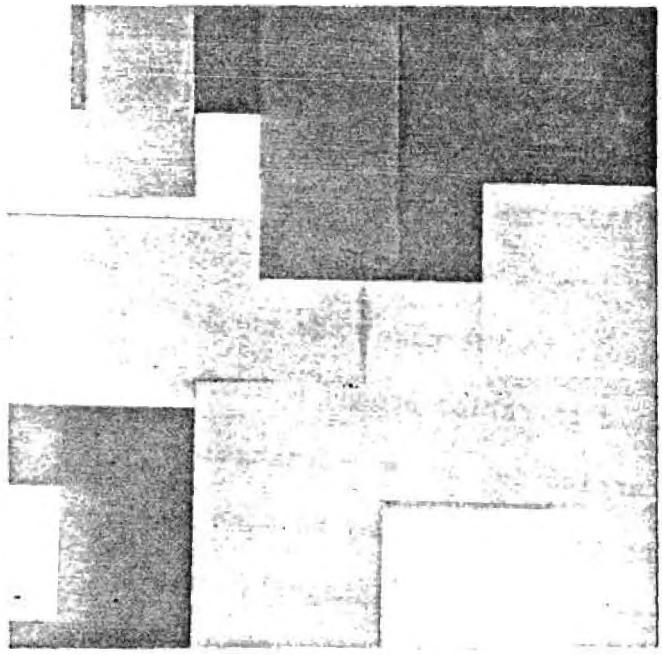


Figure 43 -- Test pattern photographed in white light

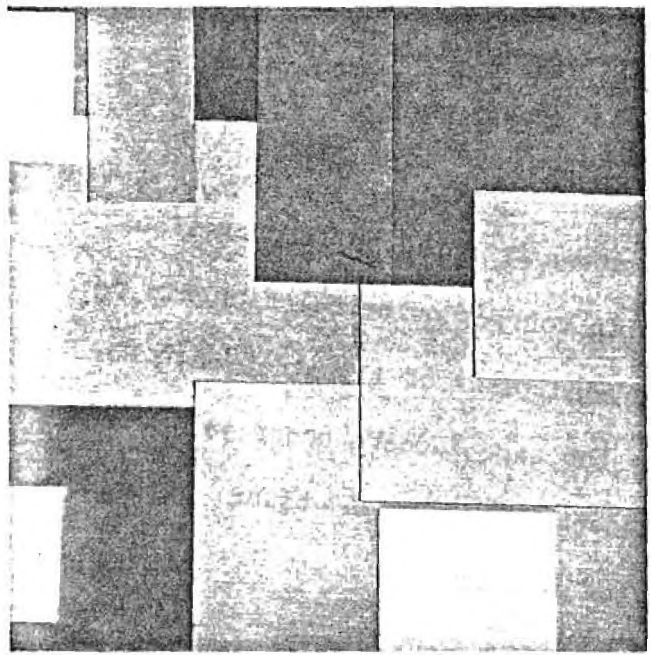


Figure 44 -- Test pattern photographed in colored light

The retinex allows for this type of change in illumination and it can also adjust for gradual shifts in illumination hue across an image.

To test these ideas a board covered with patches of colored paper was photographed in light from three slide projectors, each one fitted with either a red, green, or blue filter. The red projector, located upward and to the left of the board, supplied about twice as much light to the upper left corner as it did to the lower right corner, the green projector was positioned to supply about twice as much light to the lower right as to the upper left, and the blue projector illuminated the board evenly. These lighting conditions resulted in the pronounced red-green shift of figure 44. Note how one of the white patches of figure 43 appears pink and the other seems green. The visual system does not correct an illumination shift of this type on a reflection print as well as it would on a projected image in a darkened room. The reason is that clues from the book, paper etc. tend to inhibit the process. The red, green and blue components of figure 44 were digitized separately and each one was passed through the retinex. In figure 46 the lightness information is displayed. Note its similarity to figure 43.

Natural scenes often consist of highly textured areas rather than the cartoon style patches prepared for this experiment. The retinex does not work well on texture and this is a serious defect. The reason is that individual patches in a textured image may be small enough to occupy a single picture element, and for the threshold to work properly edges located by the gradient must not be too close together.

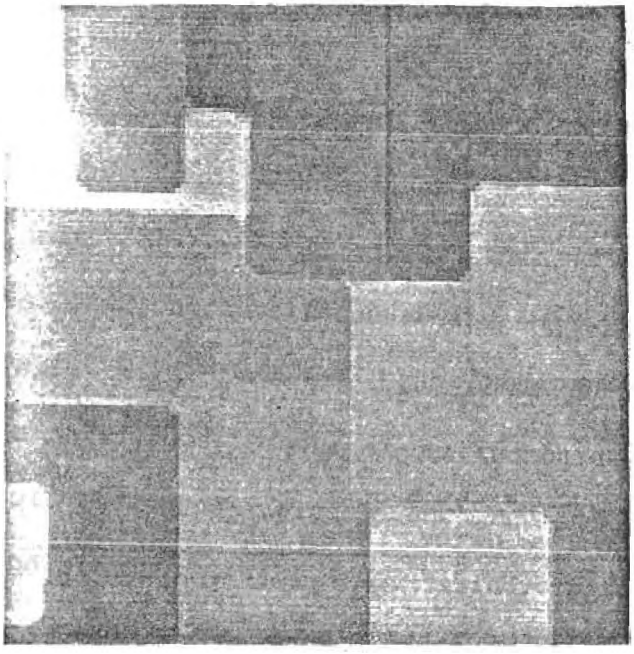


Figure 45 -- Digitized version of figure 44

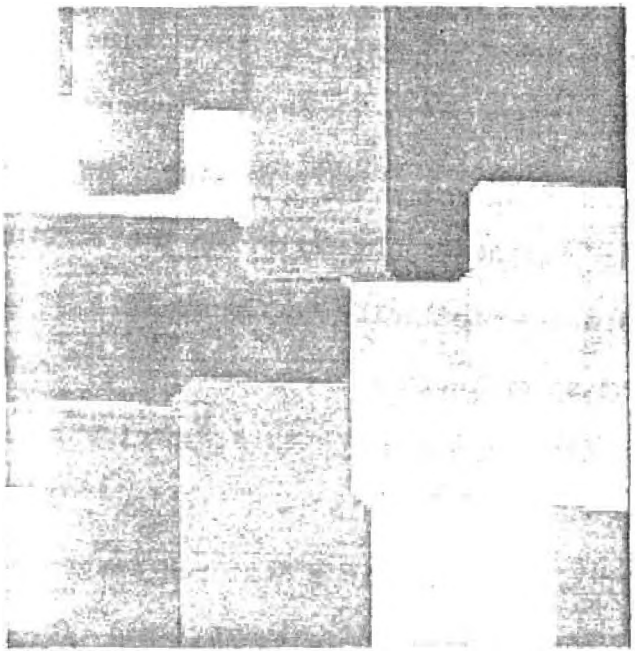


Figure 46 -- Color corrected pattern

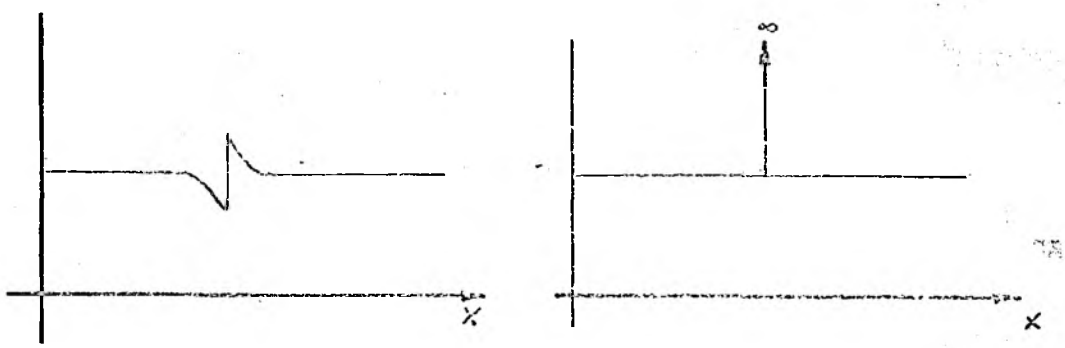
The outdoor scene in figure 15 is typical of ones with which the retinex has difficulty. In spite of this the idea is interesting and it may turn out to have further application in some other context.

4.4 The Cornsweet Illusion

Of the demonstrations presented in chapter 1, only figure 7 depends for its effect on an abrupt change in reflectance next to a gradual one. The abrupt change seems to be preserved by the visual system and the gradual one is attenuated. This property of the visual system has been taken advantage of for many years by artists and draftsmen in order to make a region in a drawing seem lighter or darker that it would otherwise. Ratliff [10] gives an interesting review of this practice. By a coincidence this is the only illusion of the collection which is predicted correctly by the retinex and ignored by the frequency selective model. Figure 47 contains plots showing how lightness information may be extracted from the images in figure 7. Note that the gradual change in intensity has been completely eliminated while the abrupt changes are preserved.

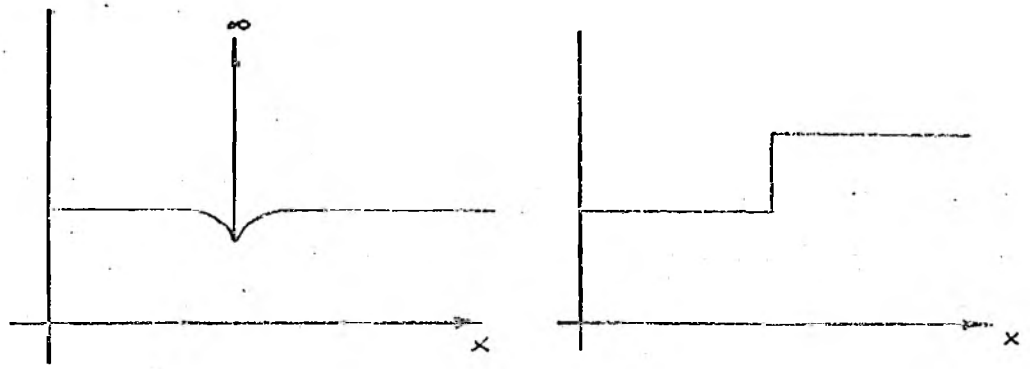
4.5 Summary

The retinex model proposes to explain how lightness information may be extracted from images by computing ratios of reflected light intensity at adjacent points and retaining ones near the boundaries of objects. Neurophysiological studies have failed to reveal an organization of the type proposed by Land, but edge information is known to be important to perception [28].



(a) Intensity function of of figure (4a)

(c) Edge information



(b) Gradient in the horizontal direction

(d) Lightness information

Figure 47 -- Extraction of lightness information from figure 7 .

The color constancy experiment of section 4.3 shows how the retinex is capable of removing illumination gradients and figure 47 shows retinex's correct response to the stimulus which produces the Cornsweet illusion.

CHAPTER 5

CONCLUSIONS

5.1 Review

The new frequency selective model of brightness perception described in chapter 3 has been shown to give subjectively useful information greater impact where it is conveyed by weak texture. Properties similar to those of the multiplicative model are retained which help subdue undesirable illumination effects, and the spatial filters acting in concert with a bank of AGC elements help overcome a limited degree of blurring. The model is the basis for a significantly improved method for restoring blurred photographs that gives results free from echo-like artifacts common to earlier methods. Restorations of images blurred in the field by both an out of focus lens and by camera motion were presented to demonstrate the method. In chapter 4 a new way of implementing the retinex was illustrated using a color constancy experiment as an example.

5.2 Speculation About Future Research

Physiological evidence is quite convincing that certain neurons in the visual cortex respond to different parts of the two dimensional Fourier transform of an image. This was illustrated by the adaptation experiment. The AGC elements are an attempt to model

the adaptation part of the demonstration, however they were designed to make the image enhancement work properly. This raises the question of whether they could be calibrated by an experiment similar to the one used by Baudelaire [11] in calibrating the multiplicative model. Another question which suggests itself from the appearance of figure 17 is whether this enhancement process could be used to make images more immune to the corrupting effects of coding. The multiplicative model has been used successfully for this purpose [20,21] and the AGC processing should bring about a further improvement. This would be a step forward with regard to finding a more accurate measure of image distortion.

Many questions about the image restoration process remain unanswered. One of them is the relationship between the cutoff frequency of the filter (figure 18) and the quality of the resotration. Circularly symmetric filters were used in the restorations reported here but is seems possible that improved results would be obtained if the symmetry depended on the type of blur (motion, out of focus etc.). Also the preparation of equalization signals should be investigated further.

APPENDIX A

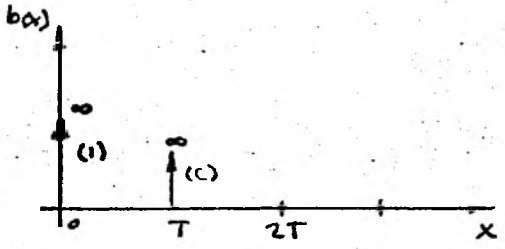
Notes on the impulse response length of convolutional inverses for space limited blurs.

Deconvolution is the process of separating signals that have been combined by convolution and it is ordinarily done with the aid of the Fourier transform.

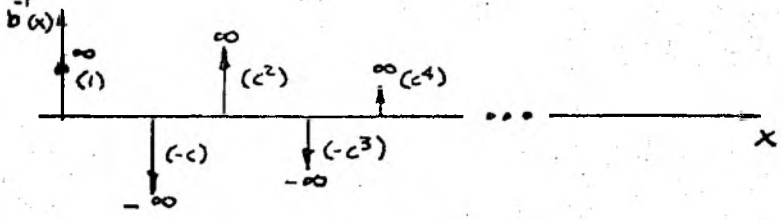
$$a = (a \circledast b) \circledast b^{-1} \xleftrightarrow{FT} (A * B)1/B$$

Suppose a signal is added to a delayed and scaled copy of itself. A convolutional representation of this situation is,

$$a \circledast b = a(t) + c*a(t - T)$$



The convolutional inverse of b can be found by inspection and is an infinite train of impulses.



Note that for c=1 the impulses do not tend toward zero for

large X.

A motion blur impluse response has a sampled representation much like the case for c=1 and its discrete convolutional inverse is of infinite extent. Truncating such a sequence will introduce errors in the deconvolution.

APPENDIX B

Notes on truncation of discrete impulse response filters

Suppose two sequences, one of which is zero outside a finite interval, are to be convolved aperiodically.

$$g(j) = \sum_{i=-\infty}^{\infty} f(i)*h(j-i)$$

where, $f(i) = 0$ for $i<k$ and $i>l$, $k<l$

If h(j) is nonzero n over the infinite interval $\infty<j<\infty$ the sum above must be taken over all i and g(j) will also be nonzero over the infinite interval.

Computational difficulties implied by the infinite sum may be avoided if g(j) is only required over part of its domain as in image processing where the output image is often made equal in size to the input image. If g(j) must be computed only for $m<=j<=n$, then

$$g(j) = \sum_{i=k}^{l} f(i)*h(j-i) \qquad m<=j<=n$$

l-k+1 terms are required in the sum, the sum must be evaluated n-m+1 times and h(j) must be available for $m-l<=j<=n-k$.

If the lengths of input and output sequences are l-k+1 and n-m+1, the impulse response must be m-l+n-k+1 points long. For example, if input and output sequences are 256 points long, 511 points of the infinite impulse response, h(j), will be required. Similar arguments apply in two dimensions.

---

* By the expression nonzero over an interval it is meant that elements of the sequence may be nonzero within the interval but they are zero outstde the interval.

APPENDIX C

Circular deconvolution of periodic sequences by filtering their log spectra

Periodic sequences combined by circular convolution may be separated using the complex logarithm * and the convolution property of the discrete Fourier transform (DFT). If a,b are sequences having the same period and A,B are their discrete Fourier transforms, then

$$a = (a \circledast b) \circledast b^{-1} \overset{FT}{\longleftrightarrow} (A * B)/B \overset{Log}{\longleftrightarrow} (\overline{A} + \overline{B}) - \overline{B} = \overline{A}$$

The square overbracket indicates the complex logarithm.

B may be removed by direct subtraction if it is known exactly or by linear filtering if its DFT is disjoint from the DFT of A.

APPENDIX D

An Image Processing Language (IPL)

Experiments of the type described herein were greatly facilitated by the development of a programming language [29] whose data types are images and whose primitives operate on images. Typical IPL primitives allow manipulation of images through addition, the discrete Fourier transform and magnitude-phase computations to name but a few. As a further experimental convenience, these operations may be invoked separately as a command language by typing interactively on a computer terminal or they may be combined into an IPL program and interpreted by the language processor. The convenience provided by this approach makes it feasible to explore many more possibilities than if each experiment required that a separate program be written, compiled and debugged using traditional editors, compilers and debugging aids.

REFERENCES

[1] T.G. Stockham Jr., "Intra-frame encoding of monochrome images by means of a phychophysical model based on non-linear filtering of multiplied signals," Picture Bandwidth Compression edited by T.S. Huang and O.J. Tretiak (New York: Gordon and Breach, 1972): pp 415-442

[2] A.V. Oppenheim, R.W. Schafer and T.G. Stockham, Jr., "Non-Linear Filtering of Multiplied and Convolved Signals," Proceedings of the IEEE 56, (August 1968): pp 1264-1291

[3] C. Blakemore and F.W. Campbell, "On the Existence of Neurons Selectively Sensitive to the Orientation and Size of Retinal Images," Journal of Physiology (London) 203, (1969): pp 237-260

[4] T.M. Cannon, "Digital Image Deblurring by Nonlinear Homomorphic Filtering," Ph.D Dissertation, University of Utah (1974)

[5] E.R. Cole, "The Removal of Unknown Blurs by Homomorphic Filtering" Ph.D Dissertation, University of Utah (1973)

[6] E.H. Land, "The Retinex," American Scientist 62 (1964): pp 247-264

[7] E.H. Land and J.J. McCann, "Lightness and the Retinex Theory," Journal of the Optical Society of America 61 No. 1, (January 1971): pp 1-11

[8] H.K.Hartline and F. Ratliff, "Visual Receptors and Retinal Interaction," Science 164, (April 1969): pp 270-278

[9] Miller, H.K. Hartline and F. Ratliff, "How Cells Receive Stimuli," Scientific American 205 No. 3, (Sept 1968): pp 222-238

[10] F. Ratliff, Mach bands: qualitative studies on neural networks in the retina, (San Francisco: Holden-Day, (1965): p 273

[11] P. Colas-Baudelaire, "Digital picture processing and psychophysics: a study in brightness perception," Ph.D dissertation, University of Utah, (1973)

[12] F.W. Campbell and J.G. Robson, "Application of Fourier Analysis to the Visibility of Gratings," Journal of Physiology (London) 197, (1968): p 551

[13] F.W. Campbell and E.R. Howell, "Monocular Alternation: a Method for the Investigation of Pattern Vision," Journal of Physiology (London) 225, (Sept 1972): pp 19p-20p

[14] T.N. Cornsweet, Visual Perception, (New York: Academic Press, 1970): p 273

[15] T.G. Stockham Jr., "Image Processing in the Context of a Visual Model," Proceedings of the IEEE 60, (July 1972): pp 828-842

[16] M.L. Davidson, "Perturbation Approach to Spatial Brightness Interaction in Human Vision," Journal of the Optical Society of America 60, (Sept 1968): pp 1300-1308

[17] G. Wyszecki and W.S.Stiles, Color Science (New York: Wylie, 1967): p 507

[18] B.R. Philip, "The Weber-Fechner Law and the Discrimination of Color Mass," Journal of Experimental Physiology 29, (1941): pp 323-333

[19] F.W. Campbell and L.Maffei, "Electrophysical Evidence of Orientation and Size Detectors in the Human Visual System," Journal of Physiology (London) 207, (1970): pp 635-652

[20] J.L. Mannos and D.J. Sakrison, "The Effects of a Visual Fideltiy Criterion on the Encoding of Images," IEEE Transactions on Information Theory IT-20 No.4, (July 1974): pp 525-536

[21] R. Rom "Image Transmission and Coding Based on Human Vision," Ph.D dissertation, University of Utah (1975)

[22] D.H. Hubel and T.N. Wiesel, "Receptive Fields and Functional Architecture of the Monkey Striate Cortex," Journal of Physiology (London) 195, (1968): pp 215-244

[23] F.W. Campblell, F.F. Cooper and Christina Enroth-Cudgell, "The Spatial Cells of the Cat," Journal of Physiology (London) 203, (1969): p 223

[24] T.G. Stockham Jr., "The Application of Generalized Linearity to Automatic Gain Control," IEEE Transactions on Audio and Electroacoustics AU-16, (June 1968): pp 267-270

[25] B. Julesz, Foundations of Cyclopean Perception, (Chicago: University of Chicago Press, 1971): pp 68-70

[26] W. Frei, "Modeling Color Vision for Phychophysical Image Processing," USCEE Report 459, September (1973): pp 112-122

[27] C.R. Wylie, Jr. Advanced Engineering Mathematics (New York: McGraw-Hil, 1966): p 582

[28] D.N. Graham, "Image Transmission by Two-Dimensional Contour Coding," Proceedings of the IEEE 55 No. 3 (March 1967): pp 336-346

[29] B. Baxter, "IPL Manual," Computer Science Memo # 7501, University of Utah